Vysoká škola báňská – Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra telekomunikační techniky

Klasifikátor detekující útoky ve VoIP infrastruktuře

Classiffier Detecting Attacks in VoIP Infrastructure.

2014 **Jan Zavadil**

VŠB - Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra telekomunikační techniky

Zadání bakalářské práce

Student: **Jan Zavadil**

Studijní program: B2647 Informační a komunikační technologie

Studijní obor: 2601R013 Telekomunikační technika

Téma: Klasifikátor detekující útoky ve VoIP infrastruktuře
Classiffier Detecting Attacks in VoIP Infrastructure.

Zásady pro vypracování:

1. Bezpečnost VoIP a rizika dle VOIPSA.
2. Analýza technik útoků v IP telefonii pomocí honeypotů.
3. Rozpoznávání útoků v IP telefonii pomocí nástroje WEKA na reálných datech.
4. Vyhodnocení úspěšnosti klasifikátoru.

Seznam doporučené odborné literatury:

PROVOS, N., HOLZ, T. *Virtual honeypots*. Publisher: Addison-Wesley Professional; 1 edition, 440 p., 2007. ISBN 978-0321336323.
SAFARIK, J., REZAC, F., VOZNAK, M., TOMALA, K., PARTILA, P. *Automatic analysis of attack data from distributed honeypot network*. Proc. SPIE 8755, Mobile Multimedia/Image Processing, Security, and Applications 2013, Baltimore, Maryland, May 2013, DOI: 10.1117/12.2015514.

Formální náležitosti a rozsah bakalářské práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

Vedoucí bakalářské práce: **doc. Ing. Miroslav Vozňák, Ph.D.**

Datum zadání: 01.09.2013
Datum odevzdání: 07.05.2014

doc. Ing. Miroslav Vozňák, Ph.D.
vedoucí katedry

prof. RNDr. Václav Snášel, CSc.
děkan fakulty

Prohlášení studenta

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

V Ostravě dne: *6. května 2014*

Jaradil

……..……
podpis studenta

Poděkování

Rád bych poděkoval doc. Ing. Miroslavu Vozňákovi, Ph.D. za odbornou pomoc a konzultaci při vytváření této bakalářské práce.

Prohlášení zástupce spolupracující právnické nebo fyzické osoby

„Souhlasím se zveřejněním této bakalářské práce dle požadavků čl. 26, odst. 9 Studijního a zkušebního řádu pro studium v bakalářských programech VŠB-TU Ostrava.“

Abstrakt

Bakalářská práce se zabývá vytvořením klasifikátoru detekující útoky ve VoIP infrastruktuře. V první části práce je popsán teoretický rozbor bezpečnosti VoIP s možnými útoky a obranami proti nim. Druhá část je zaměřena na honeypoty, jejich druhy, rozdíly a vlastnosti. Ve třetí části se práce zabývá rozpoznáním útoků pomocí rozhodovacího stromu navrženého v Matlabu a také pomocí nástroje WEKA na datech z reálného provozu.

Klíčová slova

VoIP; VoIPSA; honeypot; WEKA; SQL; Matlab.

Abstract

Bachelor's thesis aim is to create classifier for detecting attacks in the VoIP infrastructure. In the first part of thesis, security of VoIP and potential threats with corresponding defense are discussed. Second part deals with honeypots, particularly with their attributes, types a differences among them. In the third part of thesis is discussed recognizing threats using decision tree that has been designed in Matlab and another tool called WEKA, both applied on the real traffic data.

Key words

VoIP; VoIPSA; honeypot; WEKA; SQL; Matlab.

Seznam použitých zkratek

Zkratka	Anglický význam	Český význam
ARP	Address Resolution Protocol	-
DHCP	Dynamic Host Configuration Protocol	Protokol pro dynamickou konfiguraci hostitelského zařízení
DNS	Domain Name System	Server doménových jmen
DoS	Denial of Service	Odmítnutí služby
IP	Internet Protocol	Standartní internetový protokol
MAC	Media Access Control	Jedinečný identifikátor síťového zařízení
MITM	Man in the middle	Člověk uprostřed
QoS	Quality of Service	Kvalita služeb
RTP	Real-time Transport Protocol	Protokol standardizující paketové doručování zvukových a obrazových dat
SIP	Session Initiation Protocol	Protokol pro inicializaci relací
SQL	Structured Query Language	Standardizovaný dotazovací jazyk
SRTP	Secure Real-time Transport Protocol	Zabezpečený protokol standardizující paketové doručování zvukových a obrazových dat
SSH	Secure Shell	Zabezpečený komunikační protokol
TCP	Transmission Control Protocol	Spojově orientovaná protokol pracující na transportní vrstvě
URI	Uniform Resource Identifier	Jednotný identifikátor zdroje
URL	Uniform Resource Locator	Jednotný lokátor zdroje
VoIP	Voice over Internet Protocol	Internetová telefonie
WEP	Wired Equivalent Privacy	Zabezpečení bezdrátových sítí

Obsah

Úvod ..... - 11 -
1 Bezpečnost VoIP ..... - 12 -
1.1 Získání přístupu k síti ..... - 13 -
1.1.1 ARP Poisoning ..... - 13 -
1.1.2 Ochrana proti ARP Poisoningu ..... - 13 -
1.2 Modifikace ..... - 14 -
1.2.1 Obrana před modifikací ..... - 14 -
1.3 DoS útok ..... - 14 -
1.3.1 Zahlcení voláním ..... - 15 -
1.3.2 Zahlcení UDP ..... - 15 -
1.3.3 Modifikace QoS ..... - 15 -
1.3.4 Útoky na infrastrukturu ..... - 15 -
1.3.5 Poškození packetů ..... - 15 -
1.4 Odposlech ..... - 15 -
1.4.1 Provádění odposlechů ..... - 15 -
1.4.2 Ochrana proti odposlechům ..... - 16 -
1.5 Skenování a průzkum sítě ..... - 16 -
1.5.1 Sken INVITE ..... - 16 -
1.5.2 Sken REGISTER ..... - 16 -
1.5.3 Sken OPTIONS ..... - 17 -
2 Analýza útoků pomocí Honeypot ..... - 18 -
2.1 Tipy Honeypotů ..... - 18 -
2.1.1 Honeypot Artemisa ..... - 19 -
2.1.2 Honeypot Kippo ..... - 20 -
2.1.3 Honeypot Dionaea ..... - 20 -
2.2 Fyzické a virtuální honeypoty ..... - 21 -
2.2.1 Fyzické honeypoty ..... - 21 -
2.2.2 Virtuální honeypoty ..... - 21 -
2.3 Klientské a serverové honeypoty ..... - 21 -

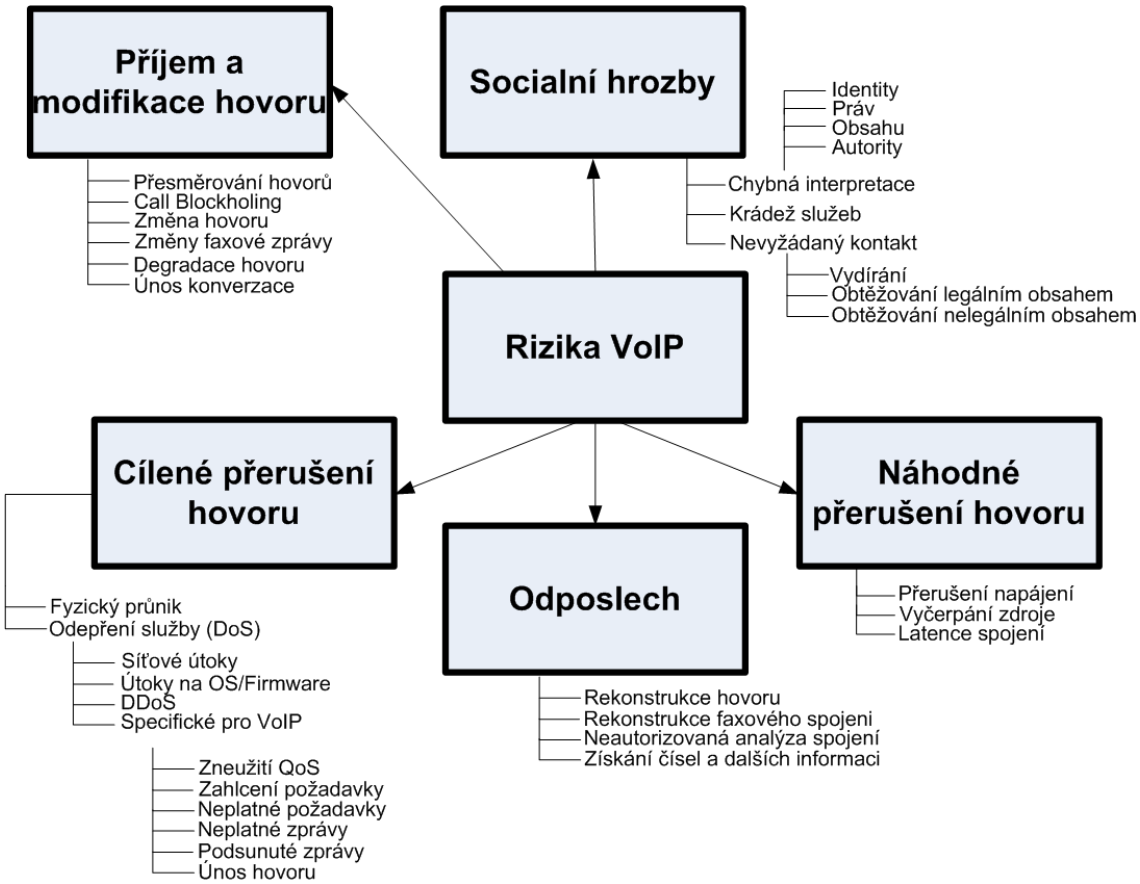
2.3.1 Klientské honeypoty - 21 -
2.3.2 Serverový honeypot - 21 -
2.4 Bezdrátový honeypot - 21 -
2.5 Honeypoty s vysokou a nízkou mírou interakcí - 21 -
2.5.1 Vysoká míra interakce - 21 -
2.5.2 Nízká míra interakce - 21 -
3 Rozpoznávání útoků praktická část - 23 -
3.1 Útoky typu INVITE - 23 -
3.2 Útok typu REGISTER - 24 -
3.3 Útok typu OPTIONS - 24 -
3.4 Návrh rozhodovacího stromu v Matlabu - 25 -
4 Klasifikace útoků pomoci nástroje WEKA - 27 -
4.1 Praktické využití nástroje WEKA - 29 -
4.1.1 INVITE - 29 -
4.1.2 REGISTER - 31 -
4.1.3 OPTIONS - 33 -
5 Vyhodnocení úspěšnosti klasifikátoru - 35 -
Závěr - 35 -
Použitá literatura - 37 -
Seznam příloh - 38 -

Úvod

Zabezpečení VoIP sítí je velmi důležité a nezbytné. Hovory totiž nemusí obsahovat jen běžnou komunikaci, ale mohou také obsahovat i citlivá data jako jsou hesla, loginy a další údaje, které může útočník využít odposlouchávání hovorů nebo jiné škodlivé činnosti. V první části jsou popsány bezpečnostní rizika z pohledu VOIPSA. Z nich jsou vypsány základní druhy útoků. Ve druhé části jsou popsány technologie honeypotů ve VoIP infrastruktuře. Jejich druhy a způsob činnost ve VoIP síti. Praktická část je věnována rozpoznání útoků a vytvořeni klasifikátoru za pomoci rozhodovacího stromu vytvořeného v Matlabu a také za pomoci nástroje WEKA. Analyzovaná databáze bude z reálného provozu zachycená honeypotom Dionaea. Za pomoci SQL dotazů budou zobrazeny jednotlivé útoky (INVITE, REGISTER, OPTIONS) rozděleny do skupin a vy následně vyhodnoceny. V poslední části budou zahrnuty výsledky všech provedených útoku.

1 Bezpečnost VoIP

Zabezpečení bezpečnosti VoIP není nijak banální záležitostí. Nové hrozby potencionálních útoků vznikají téměř s každou inovací služeb. Při tvorbě protokolů je vždy nutné myslet na mechanizmy, které zajišťují bezpečnost slabých míst proti neoprávněným přístupům. Také musíme dbát na uplatnění při implementaci. Řada VoIP protokolů vychází z emailových a webových protokolů (např. SIP), které se používají již delší dobu. Zabezpečení těchto důvěryhodných emailů a obsahu webových stránek se stává nutností alespoň v některých důležitých oblastech. Aby bylo možné provádět útoky musí být zajištěný dostatečný přístup k síti. Útok muže být založen na několika různých principech, záleží na tom, jestli útočník chce data modifikovat, nebo pouze odposlouchávat. Klasifikací a popisů různých druhů VoIP útoku je mnoho, jedním z těchto přehledů je dokument Threat Taxonomy od organizace VOIPSA. Ten popisuje více používané útoky proti VoIP sítím, viz Obrázek 1.1 [1][10].



Obrázek 1.1: *Klasifikace hrozeb a útoku dle VOIPSA*

1.1 Získání přístupu k síti

Aby útočník mohl hovory odposlouchávat, nebo provádět modifikace, musí nejdříve získat přístup k síti, kde jsou data posílána. To znamená, že se musí dostat až na úroveň paketů.

Nejjednodušší způsob pro přístup do nepřepínané sítě je použitím hubu. Hub pracuje tak, že automaticky přeposílá dál všechna data na všechny porty, bez ohledu na to, jestli data patří příslušnému uživateli. Útočník se tedy jen připojí na některý port hubu a dostane se ke všem datům, které patří účastníkovi v síti.

Pokud se používá bezdrátová Wi-Fi síť, může být útok na ni velmi jednoduchý, záleží na její konfiguraci. Pokud je Wi-Fi síť nezabezpečena je přístup do ní jednoduchý, ale i používaní zabezpečení WEP je překonatelné. V současnosti se využívá WPA šifrování.

Těžší na provedení útoku je využití switche nebo proxy serveru. Při tomto útoku záleží na konfiguraci a typu zařízení a veškeré útoky musí být řešeny individuálně. Pokud se switch konfiguruje pomocí webového rozhraní, může útočník vyzkoušet prolomení webového rozhraní a nastavit přeposílání dat na určený port. Zabezpečení proxy serveru je složitější, záleží na použití operačního systému, jeho aktuálnosti a komplexního zabezpečení. Další možností útoku na switch je záplava MAC adres. Každý switch má MAC/ARP tabulku, do které jsou ukládány informace o IP adresách. Když dojde k přeplnění kapacity tabulky, tak dojde k přepnutí switche do nouzového režimu a ten začne fungovat jako hub. Odposlech z hubu je již jednoduchý. Z toho vyplývá, že útočník může zkusit přepnout switch do nouzového režimu pomocí generování MAC adres a přeplnění MAC/ARP tabulky [2].

1.1.1 ARP Poisoning

Další možností přístupu k datům je ARP Poisoning. Útok je typu MIMT, ten je pro klienta obtížně detekovatelný. Útočník využívá toho, že některé systémy přijmou ARP zprávu a přiřadí ji do své MAC/ARP tabulky. Útočník při realizaci útoku musí zjistit MAC adresy klientů, které chce odposlouchávat. Potom pošle oběma klientům ARP zprávu o změně MAC adresy na adresu útočníka. Aby nebylo možné útočníka rozpoznat je třeba zajistit, aby se zprávy přeposílali správnému klientovi. Útočník se tedy stává prostředníkem mezi klienty, a proto má přístup k datům, které si klienti mezi sebou posílají. Tento způsob přístupu k síti dává útočníkovi možnost nejen data odposlouchávat, ale i modifikovat [2].

1.1.2 Ochrana proti ARP Poisoningu

Za nejjistější ochranu proti ARP Poisoningu je udržování statické MAC/ARP tabulky u všech důležitých síťových prvků (VoIP Proxy server, DHCP server). To muže být náročné pokud máme dynamicky vyvíjející se sítě. Další možnost máme pokud nastavíme MAC adresy pouze jednotlivým portům switche. To nám ovšem omezí používaní portu pouze na jedno zařízení, to jde ovšem obejít naklonováním MAC adresy na pomocné zařízení. Další možností ochrany je vytvoření VLAN sítě, tím můžeme oddělit VoIP data od dat, která probíhají v běžném síťovém provozu. Rozhodně se nám ale vyplatí, pokud budeme citlivá data šifrovat, jak jen to bude možné [2].

1.2 Modifikace

Modifikace zahrnuje velkou řadu útoků různých typů. Zejména se to pak týká signalizačních dat. Tyto data se snaží útočník pozměnit a tím ovlivnit spojení.

Nejtypičtějším modelem útoku je MITM (man-in-the-middle). Útočník odchytí cizí data a pozastaví je, provede jejich pozměnění a posílá je dál původnímu příjemci. Pokud je útok úspěšný, nemá příjemce žádnou možnost zjistit porušení integrity a přijme data bez jakéhokoli podezření, že byla data pozměněna. Útočník se může pokusit vydávat za někoho jiného (impersonation), může přesměrovávat další hovory (redirection), také je může ukončovat, nebo způsobovat odepření služeb (DoS). MITM útok je nutné ovlivňovat topologií sítě (např. ARP Poisoning).

Ve VoIP sítích, které jsou postaveny na SIP protokolu je možnost ovlivňování dění v síti tím, že se posílají upravené textové SIP zprávy. Jako příklad můžeme uvést odregistrování určitého klienta od proxy serveru (registration removal). To zamezí přijímat uživateli hovory. Útočník se může také pokusit o zaregistrování na určitého klienta dalším terminálem (registration addition, SIP nám dává možnost udělat registraci více terminálů na jednom klientu). Další možnost je, že se útočník pokusí přímo zaregistrovat na identitu uživatele (registration hijacking). Díky úpravám SIP zpráv může útočník také provádět vkládání nežádoucích RTP streamů, nebo úplně nahradit audio nebo video data [2].

1.2.1 Obrana před modifikací

Obrana proti modifikaci záleží na tom, jakého je útok charakteru. Nasazení TCP protokolu může útoky komplikovat. Protože TCP protokol nevyžaduje trvalé spojení jako UDP protokol. Při kontrole TCP packetů je proxy server může kontrolovat a vyřadit podvržené packety. Nasazením TLS můžeme zabezpečit další úroveň TCP.

Abychom předešli nežádoucí manipulaci, můžeme dosáhnout implementací autentizace. Při použití SIP by měly být silnými hesly autentizovány veškeré důležité zprávy (INVITE, REGISTR, BYE, atd.). Pokud je klient kompromitován, šlo by jeho registraci obnovit častým zasíláním registračních zpráv. Napadený klient se v krátkém čase přihlásí o svou identitu. Pro zajištění určité ochrany by šlo dosáhnout změnou SIP portu na jiný než 5060. Pokud nasadíme speciální SIP firewall, který by kontroloval SIP signalizaci, můžeme zajistit jistou prevenci proti útokům.

Útokům, které směřují na RTP stream je možné zabránit tím, že nastavíme šifrování a autentizaci (SRTP nebo ZRTP). Pokud RTP stream bude namixován se šifrovaným streamem, tak výsledkem po dešifrování bude nežádoucí stream znít pouze jako šum. Když máme nasazenou i autentizaci, tak útočník nemůže mixováni ani uskutečnit [2].

1.3 DoS útok

DoS útoky, nebo-li odepření služeb znemožňují, nebo komplikují, aby uživatel mohl využívat VoIP. Hlavním cílem DoS útoku není kompromitovat data, ale pouze je znepřístupnit. Typů DoS útoků je mnoho, ale všechny mají za cíl přímo VoIP. Obrana proti nim bývá složitá, a také záleží na typu útoku. Níže jsou vypsány některé z těchto útoků [2].

1.3.1 Zahlcení voláním

DoS útok je založen za posílaní velkého množství zpráv platných, nebo neplatných (např. SIP INVITE) pro sestavení hovoru na server. Také je možnost zaplavení SIP serveru pokud už proběhlo sestavení spojení, pokud už proběhla registrace, požadavky SIP INFO [2].

1.3.2 Zahlcení UDP

Při tomto útoku dojde k zahlcení a omezení přenosové kapacity spojení. Útočník jednoduše odchytí zdrojovou adresu UDP paketu, a tak zmanipuluje důvěrnou komunikaci, která probíhá mezi uživateli a projití přes zabezpečovací a filtrační systémy jako je firewall. Pro SIP protokol je tento útok obzvlášť nebezpečný, pokud je tento útok směřovaný na SIP port 5060 (standartní UDP VoIP port).

1.3.3 Modifikace QoS

Pokud dojde ke změně některého z protokolů, který není určený pro VoIP služby (např. VLAN) nebo hodnota ToS (Type of Service), muže dojít ke zpoždění paketů, nebo přijmutí v jiném pořadí [2].

1.3.4 Útoky na infrastrukturu

Jedná se o útoky, které jsou vedeny proti jednotkám infrastruktury, jako jsou DHCP server nebo DNS server, u kterých když dojde k odstavení nebo převedení do offline režimu způsobí znemožnění komunikace všech uživatelů, kteří využívají tyto služby. DNS cache poisoning je příklad jednoho z útoků, jeho funkce spočívá v oklamání DNS falešnou odpovědí, aby došlo k přesměrování bez uživatelova vědomí na falešný DNS server [2].

1.3.5 Poškození packetů

V tomto případě se vytváří různé typy paketů pro stejný protokol, který obsahuje data, která pak postupně tlačí specifikaci protokolu k bodu zlomu. Tato metoda je známá jako fuzzing. Příkladem modulu pro testování stability různých protokolů touto metodou je program ISIC. Pokud dojde odeslaní fuzzing paketu proti IP telefonu, může dojít k tomu, že telefon přestane přijímat přicházející hovory [2].

1.4 Odposlech

Odposlech (eavesdropping) je jedním z nejoblíbenějších útoků v celém VoIP. Také patří mezi jednodušší útoky. Nepatří sem jen odposlechy hlasové komunikace, ale také údaje ze signalizačních protokolů, které jsou získány neoprávněně. Aby byla zachována úplná důvěryhodnost, neměla by se dát zjistit délka hovoru, datum, ani kdo s kým mluvil.

1.4.1 Provádění odposlechů

Pokud se útočník dostane k datovému toku na úrovni jednotlivých paketů, tak za použití různých nástrojů může provádět jejich analýzu. Některé z programů dávají možnost převést RTP stream přímo na klasické zvukové formáty (wav, au, …). Útočník si je pak může přehrát na obyčejném hudebním přehrávači.

Jedním z nejznámějších programů na odchytávání komunikace je Wireshark. Používá se k analýze dat probíhajících v síti, umí také rozpoznat velké množství používaných protokolů. Používá se k analýze SIP, H.323 a dalších protokolů, ale umí také vytvořit audio soubor z RTP paketů. Další program, který umí provádět různé útoky a prolamovat hesla, je program Cain and Abel [2].

1.4.2 Ochrana proti odposlechům

Nejúčinnější ochranou je zabránění útočníkovi v připojení do sítě. U IP protokolu, ale nemáme nikdy jistotu, kudy naše data vedou, nemůžeme mít tedy nikdy jistotu, že naše data nejsou někým odposlouchávána. V současnosti je jednou z nejspolehlivějších ochran šifrování dat. Můžeme k tomu vyžít např. VPN IPsec, nebo zašifrování dat každého protokolu patřičným mechanismem. Pro přenos zvuku nebo video streamu se v dnešní době používají SRTP, popřípadě ZRTP, pro softwarové telefony Zfone.

1.5 Skenování a průzkum sítě

Jedním z prvních kroků je průzkum sítě, které musí útočník provést. Pokud útočník získá dostatek informací o infrastruktuře, naskytne se mu možnost zaměřit své útoky na slabá místa. Nemusí se jednat o nijak velkou hrozbu, detekce skenování nás může upozornit na to, že se připravuje sofistikovanější útok. Systém by měl být navržen tak, aby byl schopen odolat útoku, pokud má útočník detailnější informace a znalosti dané infrastruktury. Také si musíme dávat pozor na útoky z vnitřní sítě od uživatelů, kteří mají oprávnění.

Pro prvotní identifikaci zařízení v síti můžeme použít nmap, který je pro toto velmi vhodný. Jeho součástí je velká databáze fingerprintů VoIP zařízení. Pokud dojte k tomu, že se útočníkovi podaří získat přehled o infrastruktuře sítě, potřebuje následně zjistit uživatelská jména (tzv. SIP URI), která jsou používána. K tomu může využít následující typy skenování.

1.5.1 Sken INVITE

Provádění tohoto skenování je nejvíce nápadné. Dochází při něm k tomu, že cílový telefon vyzvání. Můžeme to považovat za nedostatek, ale na druhou stranu to můžeme použít k obtěžování uživatele.

Skenování probíhá tak, že útočník posílá požadavky o INVITE na SIP server. Na tomto serveru zkouší různá uživatelská jména, buď na základu slovníku, nebo obvyklých číselných kombinací. Pokud je uživatelské jméno platné, server nám pošle odpověď 180 Ringing. Pokud je uživatelské jméno neplatné pošle nám server zprávu 404 Not Found.

Může také nastat případ, kdy tento způsob skenování nemusí fungovat. V takovém případě server odpovídá zprávou 503 Service Unavailable. Potom lze zkusit zasílat požadavky INVITE přímo na daný telefon, a to za předpokladu, že víme jakou má IP adresu, nebo vyzkoušet jiný typ skenu.

1.5.2 Sken REGISTER

Při skenu REGISTER útočník posílá na server žádost o REGISTER s některým uživatelským jménem. Pokud server odpoví zprávou 200 OK, uživatel existuje a autentizace není vyžadována. Pokud je autentizace vyžadována, server posle zprávu 401 Unauthorized nebo

407 Prosy Authentication Reqired. Pokud uživatel neexistuje může být chování serverů různé. Některé servery budou odpovídat pořád stejně, ovšem v případě Asterisku je posílána zpráva 403 Forbidden.

1.5.3 **Sken OPTIONS**

Sken typu OPTIONS se u protokolu SIP využívá k tomu, abych zjistili jaké server, případně koncové zařízení, využívá funkce. Také může být využit na průzkum uživatelů. Tento sken je velmi nenápadný a efektivní.

Použití OPTIONS skenu probíhá podobně jako u INVITE. Útočník posílá požadavek OPTIONS na server se SIP URI (např. OPTIONS sip:607@192.168.1.50). Pokud daný uživatel existuje, zařízení odpoví zprávou 200 OK a pošle podporované parametry. Pokud neexistuje, odpoví zprávou 404 Not Found. Může také nastat situace, že skenování nebude fungovat, a v takovém případě se posílá zpráva 200 OK, i když uživatelé existují nebo neexistují.

OPTIONS sken může provádět jak vertikální tak horizontální. U vertikálního skenu je cílem obvykle SIP Proxy a využívá se na zjištění uživatelských jmen, která jsou dostupná. Horizontálním skenem se spíše zaměřujeme na rozsahy sítových IP adres, nebo na detekci živých SIP zařízení.

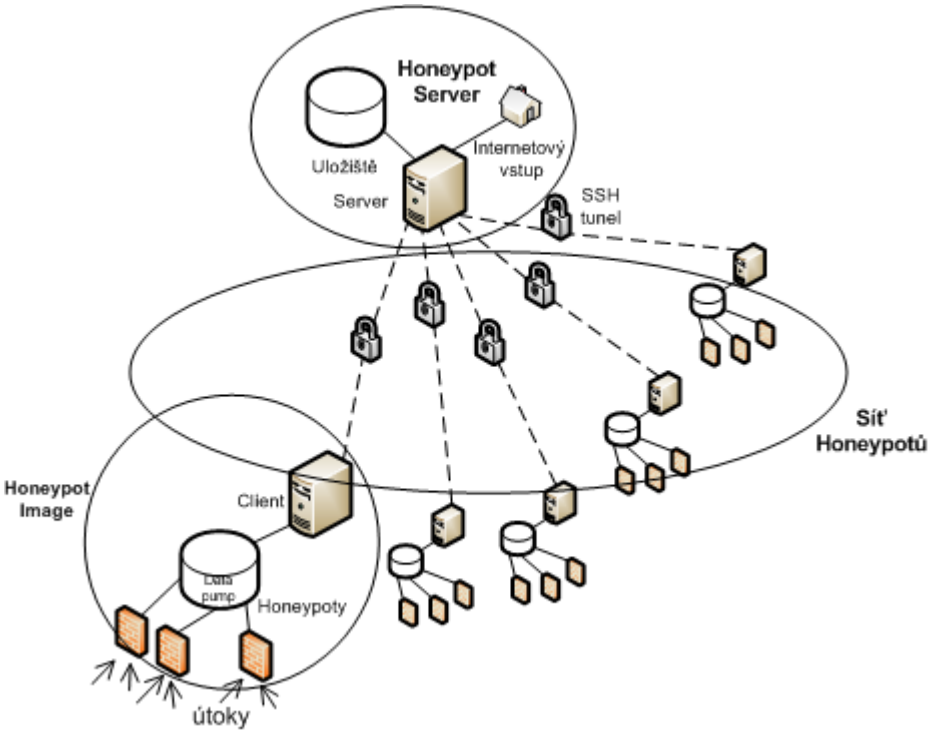
Z těchto popisů je vidět, že každý server může reagovat jinak. Ovšem obvykle je jedna z technik skenů úspěšná. Pro tento účel byla vytvořena aplikace SIPScan. Ten nám umožňuje používat všechny tři typy skenů, umožňuje i použití slovníku, který obsahuje seznam všemožných uživatelských jmen. Jako další možnost je využít linuxový balíček s nástroji, které testují VoIP SIP Vicious.

2 Analýza útoků pomocí Honeypot

Honeypot je technologie, která předvídá a zabraňuje síťovým útokům. Zabývají se sledováním a analýzou všech komunikací, které na ně směřují. Honeypot v překladu znamená hrnec medu [3], a je to systém, který láká potenciální útočníky. Můžeme si pod tím představit soubor programů, které nám v síti nevykazují žádné aktivity ani funkce. Honeypot může emulovat záměrně nebezpečnou stanici, nebo také celou síť stanic. To vyvolává u útočníka pocit, že se může dostat do systému, který obsahuje užitečná data. Veškerou komunikaci, která se takto provádí s hanypotem lze tedy považovat za neautorizovanou a škodlivou. Tuto komunikaci lze různými nástroji sledovat a analyzovat. Díky tomu můžeme pak provádět obranu proti budoucím útokům. Dalšími cíli analýzy jsou útoky, které mají za úkol využívat zranitelnosti (tzv. zero-day útok). Jedná se o zranitelnost, která nebyla dosud obecně známa nebo ještě nebyla opravená aktualizací [4].

2.1 Tipy Honeypotů

Spuštění individuální honeypot aplikace na jednom serveru přináší cenné informace. Překročení počtu spuštěných honeypotů, zejména pokud běží v různých sítích na různých geografických místech, způsobuje nežádoucí režie v analýze dat. Bez automatické agregace mechanismu může nastat situace, že údaje z honeypotů budou sníženy. Na obrázku 2.1 je znázorněné schéma s centrálním bodem pro analyzování dat [6].



Obrázek 2.1: Koncept sítě honeypot

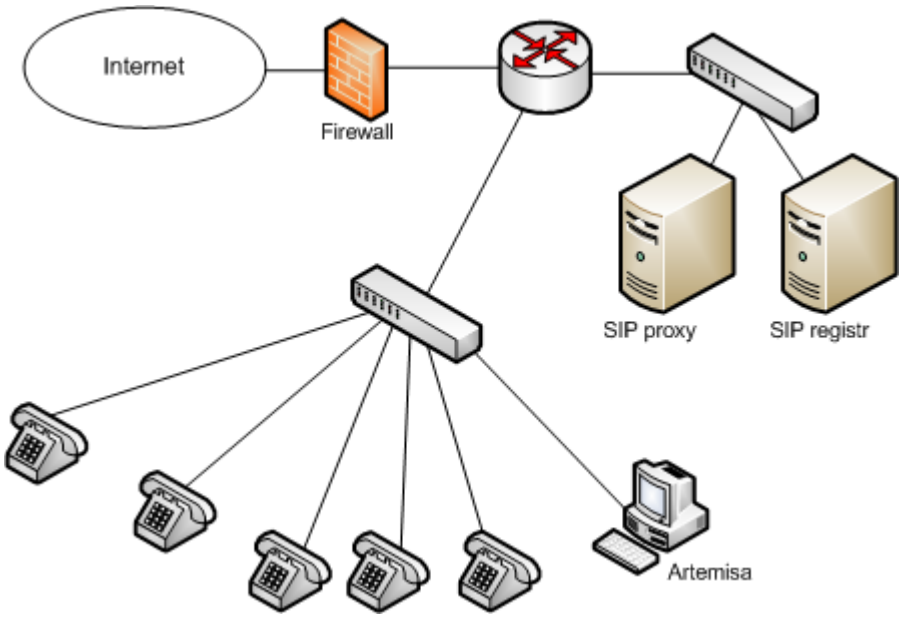
Data ze všech honeypot obrazů prochází přes datové čerpadlo a funkcí čistění. Klient pošle připravená data na server přes šifrovaný SSH tunel. Centrální server slouží jako úložiště

všech dat ze všech honeypotů, a také sleduje obrazy všech honypotů. Výsledky jsou přístupné přes webové rozhraní pro další analýzu a zlepšení bezpečnosti.

Tato architektura poskytuje snadné rozšíření pro další honeypot moduly. Honeypot Image muže obsahovat různé honeypoty nebo připojení honeypot image k jednomu serveru [6].

2.1.1 Honeypot Artemisa

Tento honeypot lze nasadit v libovolné VoIP infrastruktuře, která používá SIP protokol (Obrázek 2.2).



Obrázek 2.2: Koncept VoIP sítě s honeypotem

Program se připojí k SIP proxy s příponami, které jsou definovány v konfiguračním souboru. Rozšíření by měla být v rozsahu, která se obvykle používají pro skutečné účty. Hlavním cílem je vytvořit lepší maskování proti potenciálním útočníkům. Artemis sám nesimuluje PBX, ale je spíše aktivním koncovým zařízením [5].

Jakmile je hovor založen na jedno rozšíření Artemis, tak honeypot jednoduše odpoví na volání. Zároveň začne zkoumat příchozí SIP zprávy. Artemis pak klasifikuje volání a uloží výsledek pro další přezkoumání ze strany správce zabezpečení.

Zpráva je zařazena v následujících krocích. Za prvé, Artemis hledá otisky známých útoků. Pokud útočník používá některý z populárních útoků můžeme snadno odhadnout jeho záměry. Potom se zkontrolují doménová jména a SIP porty na straně útočníka (za předpokladu že jsou otevřeny). K dispozici jsou také podrobné kontroly mediálních portů. Požadované URI jsou takto kontrolovány stejně jako přijaté ACK zprávy od uživatele. A nakonec, Artemis kontroluje také přijatý RTP stream, za předpokladu že RTP stream byl založen (audio stopa z přijatého hovoru lze uložit do WAV formátu).

Tento sled postupů pomáhá Artemis klasifikovat volání. Výsledky se potom zobrazují v konzole. Můžeme je také uložit do předem určené složky, nebo je odeslat jako upozornění emailem [5].

2.1.2 Honeypot Kippo

Toto je další používaný honeypot, který je ovšem založen na jiných principech. Není orientován přímo na VoIP jako Artemisa. Kippo spíše simuluje SSH server. Když se útočník pokusí dostat na server je přesměrován do honeypotu. Tato situace platí, pokud útočník není na seznamu povolených IP adres. Po připojení k honeypotu dojde k jejímu založení, útočník musí zadat správné uživatelské jméno a heslo. Honeypot je nastaven na nejčastější uživatelské jméno ROOT a hesla viz. tabulka 3.1. Další kombinace se ukládají do souboru data/pass.db.

Heslo	Počet
	28146
123456	17625
password	6325
1234	5663
12345	5501
123	5342
1qa2ws3ed	5278
a	5121
test	4743
qwerty	4601

Tabulka 2.1: 10 nejčastěji požívanách hesel

Kippo zaznamená každý pokus o přihlášení. V případě, že je zadaná kombinace platná, je útočníkovi umožněn přístup k falešným souborům. Všechny záznamy, které Kippo provádí, jsou uloženy v databázi MySQL pro usnadnění následující analýzy [5].

2.1.3 Honeypot Dionaea

Dionaea je více orientovaný honeypot, který dokáže simulovat více služeb najednou. Obvykle jsou informace z těchto služeb pouze obecné, ale Dionaea pracuje jen s určitým počtem, jako je SMB (sdílení tiskáren, souboru, sériových portů), HTTP, FTP, TFTP, MSSQL, SIP protokoly. Útočníci ve většině případů zneužívají právě tyto protokoly. Dionaea má také možnost ukládat nebezpečný obsah.

Dionaea pracuje jiným způsobem než Artemis. Není zapotřebí připojení k externímu VoIP serveru. Čeká na jakékoli SIP zprávy a snaží se na ně odpovědět. Podporuje všechny požadavky SIP z RFC 3261 (REGISTER, INVITE ACK, CANCEL, BYE, OPTION). Dionaea podporuje více SIP relací a RTP audio streamy (data streamy mohou byt zaznamenány) [5].

2.2 Fyzické a virtuální honeypoty

2.2.1 Fyzické honeypoty

Toto využití honeypotu je náročné kvůli finanční zátěži, protože fyzický honeypot je realizován na hardwarovém stroji. To znamená, že jedna stanice je určena pouze pro jeden honeypot a jeho veškeré funkce a potřeby [7].

2.2.2 Virtuální honeypoty

U virtuálních honeypotů je výhodou, že díky virtualizaci můžeme provozovat najednou více operačních systémů na jednom fyzickém stroji. To znamená, že pokud použijeme více fyzických strojů můžeme pak vytvořit a provozovat síť s velkým množstvím honeypotů [7].

2.3 Klientské a serverové honeypoty

2.3.1 Klientské honeypoty

Klientský honeypot simuluje koncového uživatele. Proto se musí aktivním způsobem procházet webové stránky a vyhledávat infiltrované servery [7].

2.3.2 Serverový honeypot

Serverový honeypot je přesně opačný. Čeká dokud na něj útočník nezaútočí. Až dojde k útoku, nebo pokusu o komunikaci s touto stanicí, zahájí se okamžité sledování a analýza [7].

2.4 Bezdrátový honeypot

Bezdrátové honeypoty se od ostatních trochu liší. Jejich úkolem je chránit bezdrátové sítě. Vytváří fiktivní bezdrátové přístupové body, které mají slabé, nebo žádné zabezpečení. Toto pak poskytne možnost útočníkovi o připojení do sítě (např. pro připojení k internetu) [7].

2.5 Honeypoty s vysokou a nízkou mírou interakcí

2.5.1 Vysoká míra interakce

Tyto honeypoty mají vlastnost emulovat celý systém s velkým množstvím aplikací a služeb. Díky možnosti použití mnohem většího počtu nástrojů můžeme o útočníkově pokusu zjistit mnohem více informací. Jejich nevýhodou je velká náročnost na správu a na implementaci. Za další nevýhodu je možné považovat to, že pokud útočník ovládne celý systém, může jej pak využít k dalším útokům a škodlivé činnosti [7].

Představitelé těchto honeypotů: Cupture-HPC, Honey@Home, Argos.

2.5.2 Nízká míra interakce

Schopnost těchto honeypotů je v tom, že mohou emulovat jen určité programy, funkce nebo služby. Emulace jsou však do jisté míry omezeny. Jedná se o dané odpovědi na určité útočníkovi akce. Je to jednoduchý systém z hlediska údržby a implementace, ovšem poskytují nám velmi malé množství informaci o útocích, jedná se spíše o statistiky. Nejznámější honeypot

z této skupiny je Honeyd. Tento honeypot může vytvářet virtuální hosty na volných IP adresách a emulovat na nich služby a operační systémy [7].

Představitelé těchto honeypotů: HoneyC, Tiny Honeyport.

3 Rozpoznávání útoků praktická část

Pro rozpoznání útoků jsem dostal databázi z honeypotu Dionaea. Pro vyčlenění jednotlivých informací jsem používal SQL dotaz. Po získání dat jednotlivých útoků jsem vytvořil sjednocenou databázi, která byla zapotřebí pro rozhodovací strom, který jsem navrhl v Matlabu.

Tento příkaz vypíše SIP command methody u níž byl pokus o přípojeni více než 3.

3.1 Útoky na INVITE

Metoda INVITE slouží k přizvání uživatele nebo služby k podílení na relaci. Tělo této zprávy obsahuje popis spojení.

Pro získání dat jsem použil SQL dotaz:

```
SELECTsip_commands.sip_command_method,connections.remote_host,
connections.remote_port,connections.connection,connections.conne
ction_root, COUNT(connections.connection_root) FROM sip_commands
INNER JOIN connections ON sip_commands.connection =
connections.connection WHERE sip_commands.sip_command_method =
'INVITE' GROUP BY connections.connection_root HAVING
COUNT(connections.connection_root) > 3
```

Tabulka 3.1: Výpis SQL donazu pro INVITE

sip command method	remote host	remote port	connection	connection root	Počet útoků
INVITE	37.8.54.135	10083	382	376	6
INVITE	37.8.54.135	10083	408	383	25
INVITE	37.8.54.135	10083	419	409	10
INVITE	87.230.26.233	5060	935	906	29
INVITE	198.101.195.124	5082	958	957	1
INVITE	198.101.195.124	5088	960	959	1
INVITE	198.101.195.124	5082	962	961	1
INVITE	198.101.195.124	5082	964	963	1
INVITE	198.101.195.124	5071	968	966	2
INVITE	198.101.195.124	5078	970	969	1

Celá tabulka viz Příloha A.

3.2 Útok na REGISTER

Metoda REGISTER se používá k zaregistrování současné adresy klienta na SIP serveru, tento server ji předává lokalizační službě.

Pro získání dat jsem použil SQL dotaz:

```
SELECT sip_commands.sip_command_method, connections.remote_host,
connections.remote_port, connections.connection,
connections.connection_root, COUNT(connections.connection_root)
FROM sip_commands INNER JOIN connections ON
sip_commands.connection = connections.connection WHERE
sip_commands.sip_command_method = 'REGISTER' GROUP BY
connections.connection_root HAVING
COUNT(connections.connection_root) > 3
```

Tabulka 3.2: Výpis SQL dotazu na REGISTER

sip command method	remote host	remote port	connection	connection root	Počet útoků
REGISTER	64.31.62.123	5189	345	345	16
REGISTER	37.8.54.135	10083	376	376	1
REGISTER	37.8.54.135	10083	420	420	1
REGISTER	218.95.228.109	5080	448	448	190
REGISTER	218.95.228.109	5088	449	449	75
REGISTER	64.31.62.120	5918	628	628	1
REGISTER	64.31.62.120	5528	629	629	45
REGISTER	109.169.37.147	5075	681	681	70
REGISTER	109.169.37.147	5077	682	682	694
REGISTER	125.39.0.37	5100	736	736	225

Celá tabulka viz Příloha A.

3.3 Útok na OPTIONS

Metoda OPTIONS žádá a zjišťuje informace o vlastnostech účastníka, bez toho, aby bylo vytvořeno spojení.

Pro získání dat jsem použil SQL dotaz:

```
SELECT sip_commands.sip_command_method, connections.remote_host,
connections.remote_port,connections.connection,
connections.connection_root,COUNT(connections.connection_root)
FROM sip_commands INNER JOIN connections ON
sip_commands.connection = connections.connection WHERE
sip_commands.sip_command_method = 'OPTIONS' GROUP BY
connections.connection_root HAVING
COUNT(connections.connection_root) > 0
```

Tabulka 3.3: Výpis SQL dotazu na OPTIONS

sip command method	remote host	remote port	connection	connection root	Počet útoků
OPTIONS	199.195.212.40	5066	365	365	1
OPTIONS	199.195.212.32	5063	425	425	1
OPTIONS	213.221.25.68	5065	434	434	1
OPTIONS	108.59.6.3	5078	436	436	1
OPTIONS	85.195.82.165	5060	443	443	1
OPTIONS	67.18.220.50	5060	445	445	1
OPTIONS	218.95.228.109	5071	447	447	1
OPTIONS	64.27.3.24	5088	453	453	1
OPTIONS	114.247.18.4	5060	454	454	1
OPTIONS	199.195.212.32	5130	456	456	1

U tohoto typu útoku dochází k tomu, že jedna IP adresa zkouší pokaždé jiný port. To znamená, že ve výpisu jsou všechny brány jako jediné připojení. Proto jsem použil SQL dotaz, který vypíše útočící IP adresu a počet jejich připojení. Celá tabulka viz Příloha A.

```
SELECT connections.remote_host, COUNT(*) FROM connections INNER
JOIN sip_commands ON connections.connection =
sip_commands.connection WHERE sip_commands.sip_command_method =
'OPTIONS' GROUP BY remote_host;
```

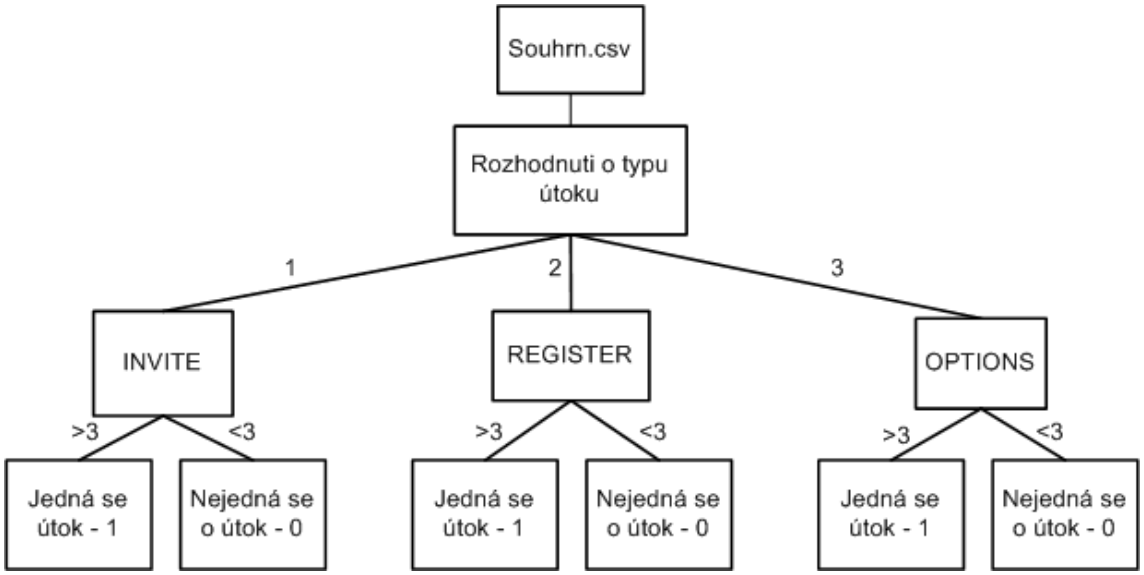
Tabulka 3.4: Výpis SQL dotazu na IP OPTIONS

sip command method	remote host	Počet útoků
OPTIONS	199.195.212.31	1
OPTIONS	199.195.212.32	19
OPTIONS	199.195.212.40	9
OPTIONS	202.103.52.147	2
OPTIONS	207.36.27.127	3
OPTIONS	208.115.236.36	1

Celá tabulka viz Příloha A.

3.4 Návrh rozhodovacího stromu v Matlabu

V programu Matlab jsem napsal rozhodovací strom, který po vložení dat rozpozná a rozhodne, zda se jedná o útok nebo ne. Data jsou ve formátu Souhrn.csv. Protože Matlab načítá pouze numerické hodnoty, byly přejmenovány názvy útoku na čísla INVITE - 1, REGISTER - 2 a OPTION - 3. Určení útoku je pomocí 0 (nejedná se o útok) nebo 1 (jedná se o útok) obrázek 3.1. Program tyto útoky vypíše a označí je spojením (platí pro INVITE a REGISTER útok) popřípadě IP adresou (pokud se jedna o útok OPTIONS). Program v Příloze B.



Obrázek 3.1: Rozhodovací strom pro Matlab

Tabulka 3.5: Výsledky z rozhodovacího stromu

INVITE:		**REGISTER:**				**OPTIONS:**			
Výsledek	**Spojení**	**Výsledek**	**Spojení**	**Výsledek**	**Spojení**	**Výsledek**	**IP adresa**	**Výsledek**	**IP adresa**
0	2	0	1	1	629	0	108.171.179.137	0	207.36.27.127
0	40	0	53	1	681	0	108.171.179.151	0	208.115.236.36
1	382	0	54	1	682	0	108.171.189.124	0	213.221.25.68
1	408	0	57	1	736	0	108.59.6.3	0	216.14.120.137
1	419	0	101	0	737	0	109.169.37.147	0	216.187.153.142
1	935	0	103	1	805	0	109.169.41.173	0	218.3.160.244
0	958	0	105	1	808	0	114.247.18.4	0	218.95.228.109
0	960	0	107	1	938	0	121.241.12.134	0	219.134.60.18
0	962	0	109	1	939	0	125.39.0.37	0	5.10.78.244
0	964	0	111	1	945	0	128.204.196.26	0	5.10.78.246
0	968	0	113	0	946	0	129.176.60.34	0	64.27.3.24
0	970	0	114	1	953	0	142.4.34.25	0	64.31.45.51
0	973	0	291	1	954	0	184.106.134.105	0	64.31.62.120
0	975	0	302	1	991	0	184.82.2.66	0	64.31.62.122
0	977	1	344	1	992	0	188.40.199.196	0	64.31.62.123
0	980	1	345	0	1000	0	188.40.199.198	0	67.18.220.50
0	982	0	376	0	1001	0	188.40.199.201	0	69.162.77.51
0	984	0	420	1	1003	0	199.195.212.31	0	75.127.3.253
0	987	1	448	1	1004	1	199.195.212.32	0	85.195.82.165
0	989	1	449	1	1009	1	199.195.212.40	0	85.195.82.178
		0	628	1	1010	0	202.103.52.147	0	86.109.125.244

4 Klasifikace útoků pomoci nástroje WEKA

Tento analytický nástroj byl vyvinut na univerzitě Waikato na Novém Zélandě. Byl celý přepracován do jazyku Java, a to v roce 1997. Díky tomu je možnost snadné modifikace na většině používaných operačních systémů. Program byl původně zaměřen na analýzu v zemědělství. Program byl ovšem později upraven pro obecné použití. WEKA obsahuje nástroje, které nám umožní předzpracování dat, vizualizaci, regresi, sdružovací pravidla, klasifikaci a shlukování. Pro vstupní data můžeme použít SQL databázi (ta je použita v praktické části). V dnešní době je do systému implementována podpora výpočtů pomocí Gridu.

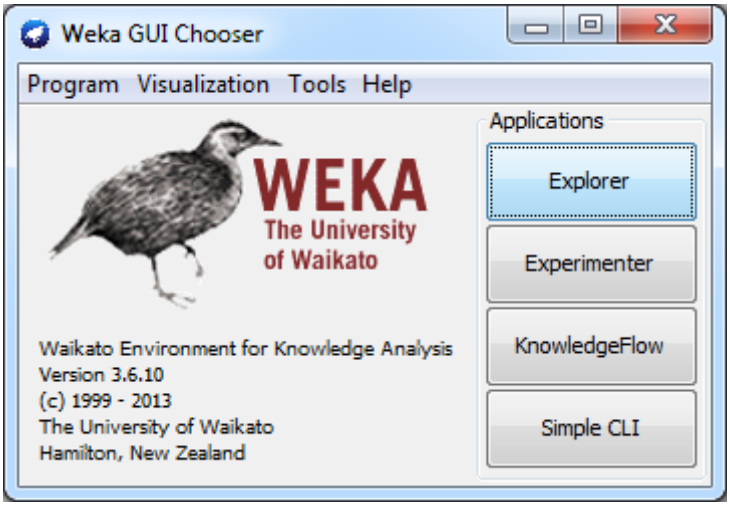
Program WEKA je open source nástroj na získávání znalostí z dat, která jsou volně dostupná. Je udržován stálými aktualizacemi. Veškeré nové algoritmy jsou po jejich objevení přidány. WEKA nabízí stovky různých algoritmů pro získávání znalostí, dalších padesát algoritmů pro předzpracování dat, a také velké možnosti vizualizace dat. Algoritmy, které WEKA poskytuje, jsou především pro klasifikaci, regresi, asociaci, shlukování a výběr atributů. Poskytuje nám možnost využít různé algoritmy a jejich srovnání za účelem zjištění, který z algoritmů je nejvhodnější pro daný problém.

WEKA je pro uživatele, kteří nejsou v oblasti získávání znalostí odborníky, relativně snadno použitelná. WEKA se snaží hlavně podporovat koncové uživatele strojového učení. Tito uživatelé mají již základní znalosti o získávání znalostí a rozumí používaným datům.

WEKA nám poskytuje několik způsobů, jak pracovat s daty. Učící metodu můžeme aplikovat na data a jak analyzovat vstup. Účelem je zjištění co nejvíce informací o datech. Modely, které se takto WEKA naučí, můžeme použít na generování dalších předpovědí na nových datech [8].

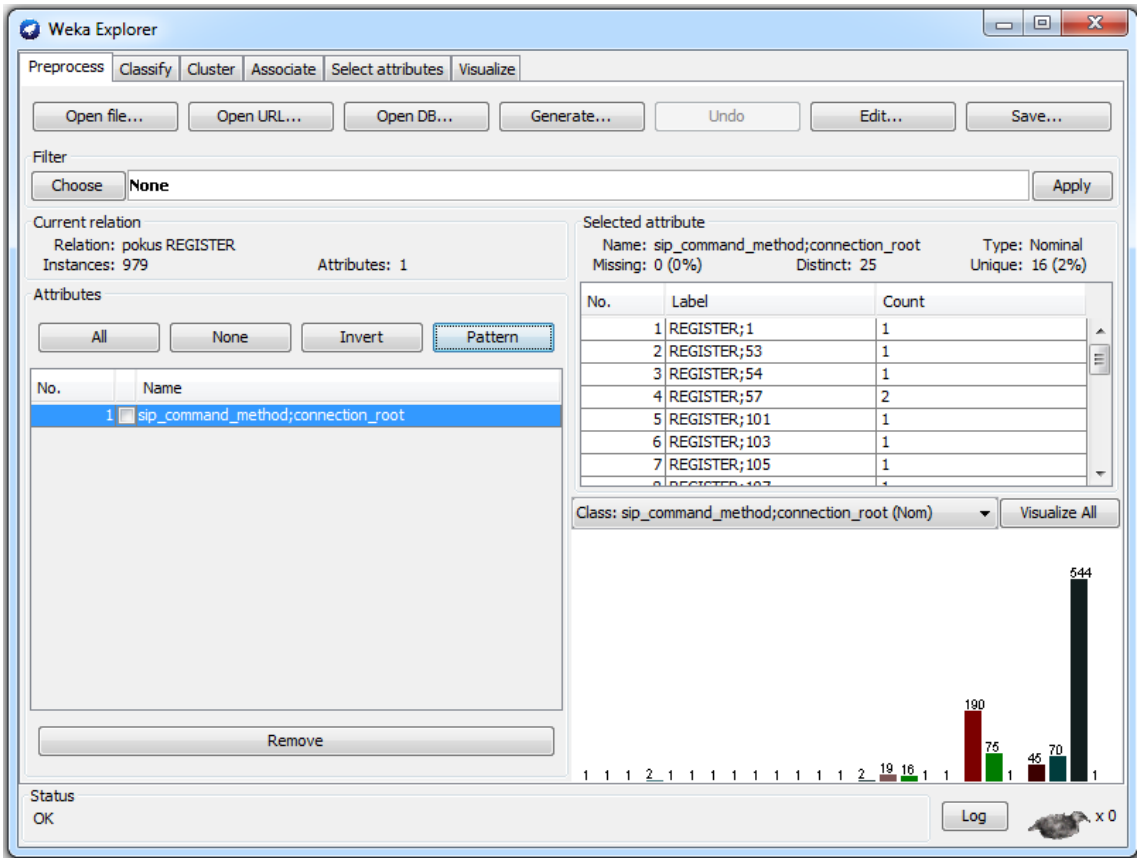
Jako vstup mohou algoritmy přijímat data ve formě relační tabulky, a to ve formátu ARFF. Data mohou být načtena ze souboru, nebo databázovým dotazem. Hlavička souboru ARFF obsahuje název relace, seznam použitých atributů a jejich použité typy a nakonec samotná data [9]. WEKA nám dává možnosti načítat data různých formátů (např. .arff, .csv). Data jsou konvertována do formátu ARFF. WEKA má také možnost načítat data ze stránek. Stačí zadat URL, ze stránek se stáhne soubor ARFF, nebo můžeme data získat pomocí SQL Select dotazů z databáze. Vstupem tedy bude relace, kterou WEKA načte jakou soubor ARFF.

Protože je program řešen jako knihovna programů, který je napsán v jazyce Java, volí si jednotlivá rozhraní. WEKA nám nabízí čtyři uživatelská rozhraní. Jsou to EXPLORER, KNOWLEDGE FLOW, EXPERIMENTER a příkazový řádek SIMPLE CLI (Obrázek 4.1). První tři nám nabízí grafická rozhraní.



Obrázek 4.1: WEKA GUI

Rozhraní Explorer je dále rozděleno do šesti panelů, každý z nich je určen pro jinou úlohu získávání dat. První panel slouží pro předzpracování dat, je v něm zahrnuta i nabídka algoritmů. Dále nám poskytuje informace o datech a také jejich grafickou prezentaci, Obrázek 4.2.



Obrázek 4.2: WEKA Explorer

Čtyři následující panely nabízí algoritmy pro získávání funkce ve stromové struktuře, a také textovou prezentaci. Poslední panel Vizualizace nám nabízí možnost dvoudimenzionální prezentace dat podle námi nastavených dvojic atributů. Toto rozhraní má ovšem tu nevýhodu, že může pracovat jen s malými nebo středně velkými množinami dat, a to z toho důvodu, že data se udržují načtená v hlavní paměti.

Rozhraní Knowledge Flow poskytuje uživateli osm panelů. Panely obsahují seznamy funkcí ve formě komponent, ty se přidávají postupně a spojují se do orientovaného grafu, jenž analyzuje a zpracovává data. Zde můžeme zpracovávat i velké datové množiny, a to za předpokladu, že jsou použité funkce, které mají možnost inkrementálního učení.

Rozhraní Experimenter umožňuje uživateli porovnávat výsledky různých algoritmů a různých datových souborů. Výsledkem toho jsou informace o nastavení parametrů a algoritmů, která jsou pro daná data nejvhodnější.

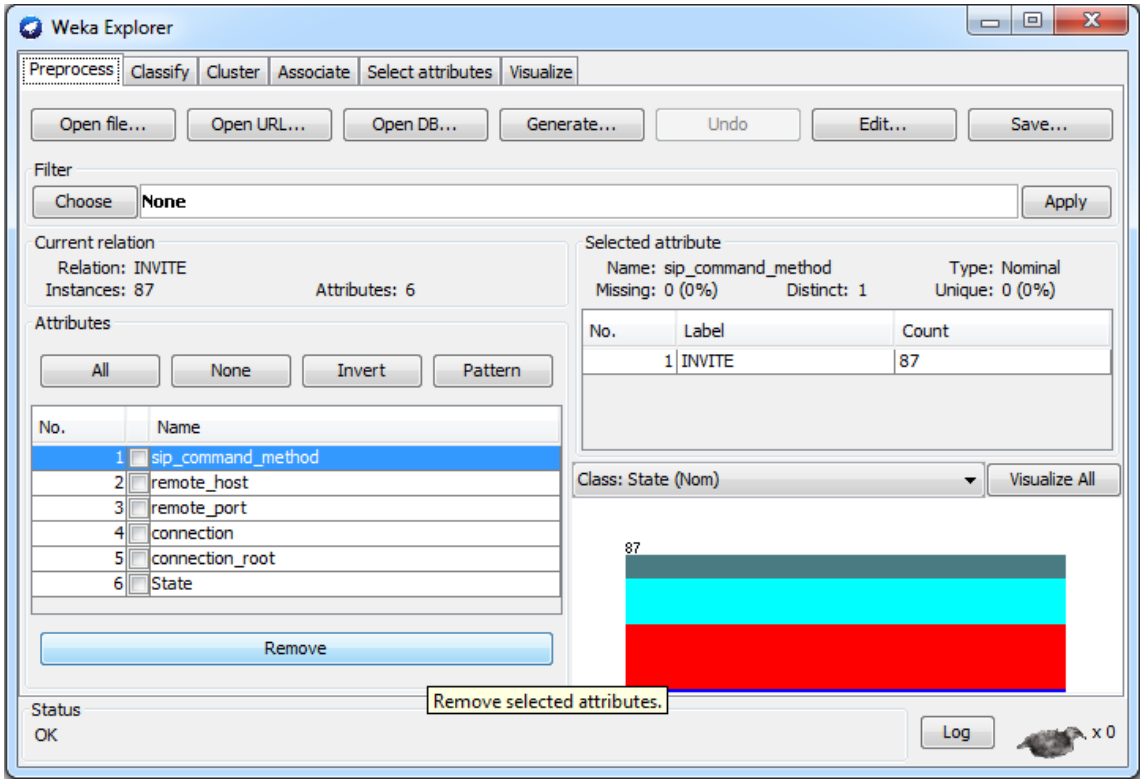
Simple CLI je příkazový řádek, který přistupuje k možnostem systému za pomocí textových příkazů

4.1 Praktické využití nástroje WEKA

Nástroj WEKA dokáže nahrát soubor ve formátu CSV, ale pro lepší a přehlednější práci jsem si CSV soubor převedl do formátu ARFF. Tento formát je přímo podporován nástrojem WEKA. Pro převedení jsem použil konvertor, který jsem našel na internetu (http://slavnik.fe.uni-lj.si/markot/csv2arff/csv2arff.php?do=home). Stačí nahrát CSV soubor, nastavit nominální hodnoty a soubor převést. Pro přehlednější práci pracuji vždy pouze s jedním typem útoku (INVITE, REGISTER, OPTIONS), ale pro ukázku jsou v Příloze C ukázány výsledky i ze všech útoku dohromady. Pro analýzu jsem dostal data z honeypotu Dionaea.

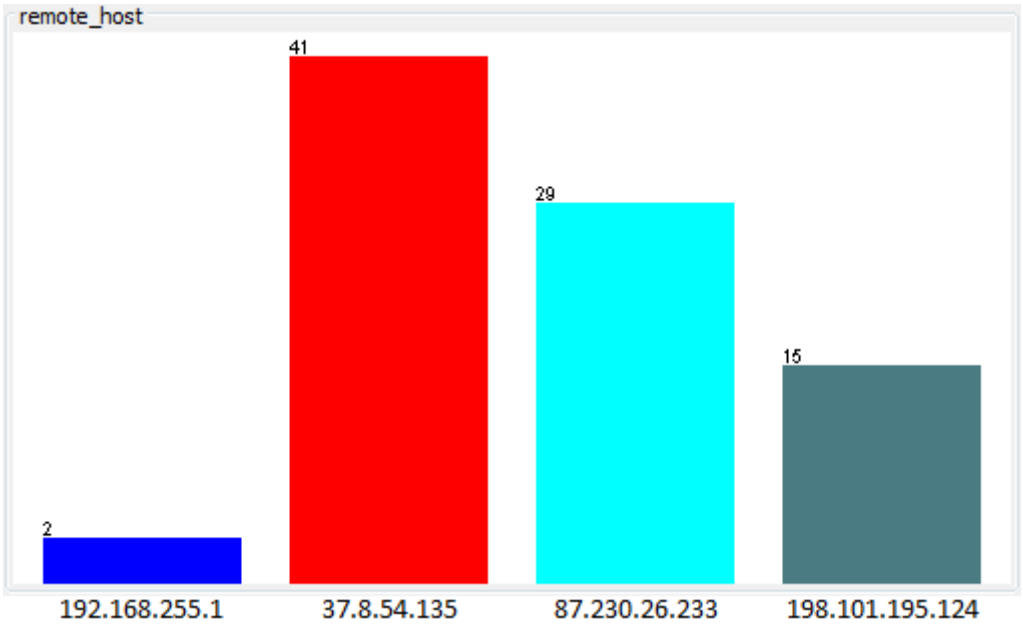
4.1.1 Útok na INVITE

Pro analýzu INVITE útoku jsem si nejdříve pomocí SQL dotazu získal z databáze všechny potřebné atributy. Nastavené atributy byly SIP_command_metod (pro určení typu útoku), remote_port (pro určení zrojové IP adresy útoku), remote_port (pro určení zdrojového portu útoku), connection (pro určení spojení), connection_root (pro určení kořene spojení) a States (pro určení země původu útoku). V každém atributu jsou obsaženy zdrojová data z databáze. Atributy jsou pro všechny analýzy útoků stejné. Po překonvertování vzniklého CSV do souboru ARFF provedla WEKA analýzu dat a připravila vizuální zpracování (viz Obrázek 4.3).



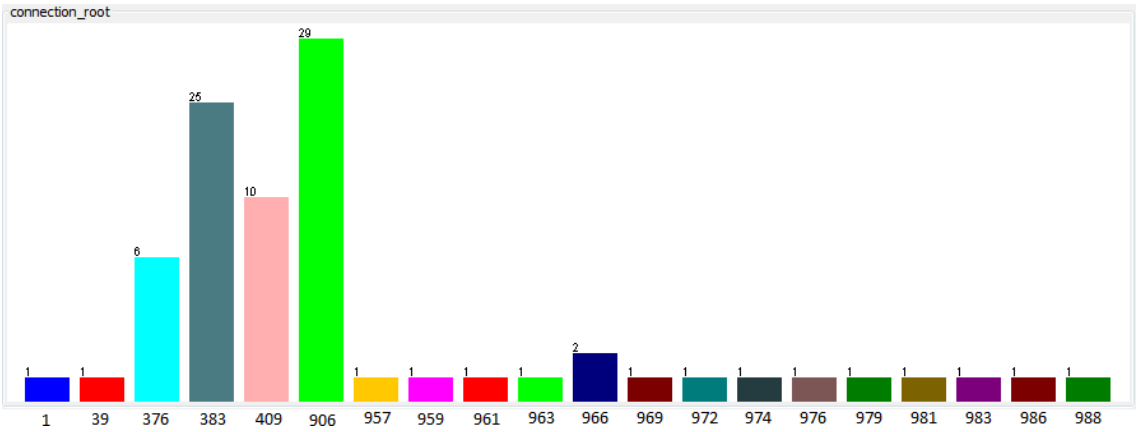
Obrázek 4.3: Atributy INVITE v nástroji WEKA

WEKA analyzovala všechna data a analyzovala veškerá data, z následujícího grafu je vidět IP adresy a jejich počet, které prováděli INVITE útok, viz Obrázek 4.4.



Obrázek 4.4: IP adresy provádějící útok na INVITE

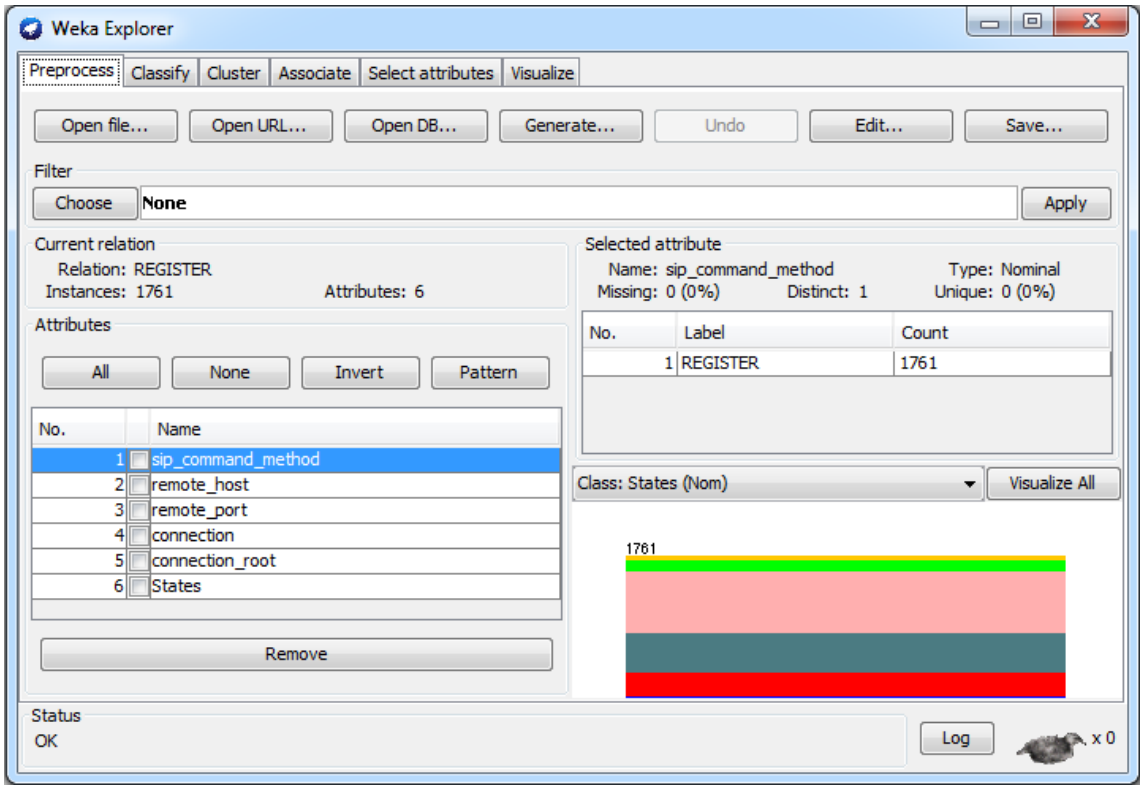
Na Obrázku 4.5 jsou zobrazena kořenová spojení a počet útoků, které byli provedeny v daném spojení.



Obrázek 4.5: Kořenová spojení a počet provedených útoků na INVITE

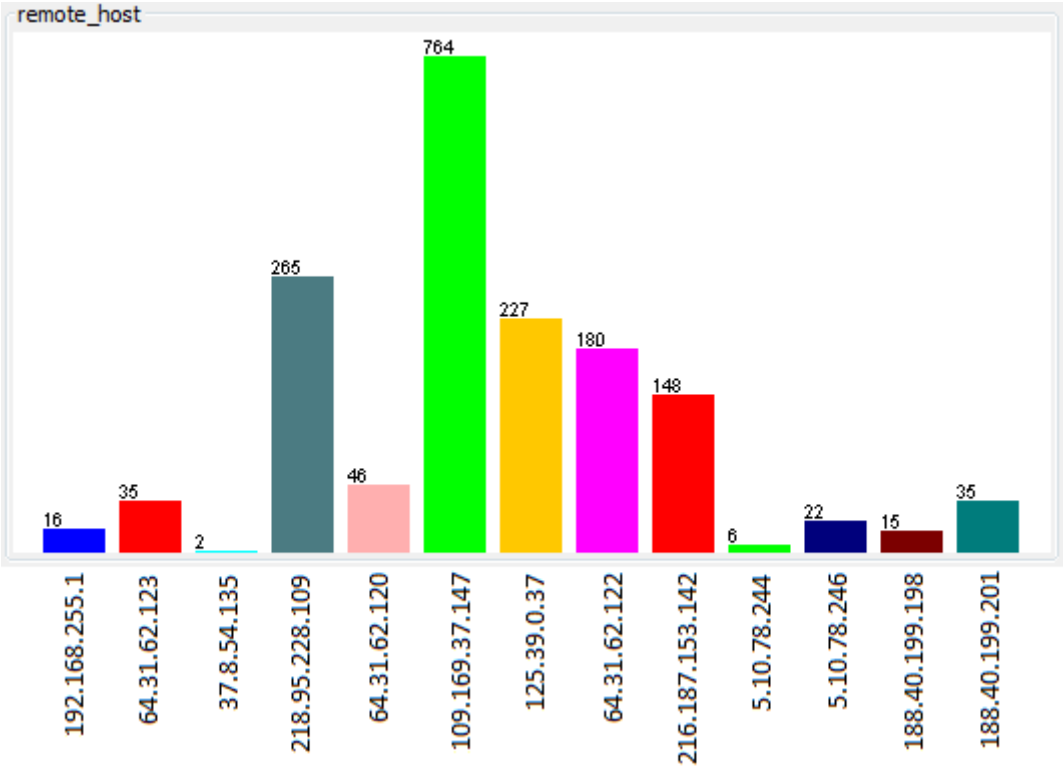
4.1.2 Útok na REGISTER

Pro analýzu REGISTER útoku jsem si nejdříve pomocí SQL dotazu získal z databáze všechny potřebné atributy. Nastavené atributy byly SIP_command_metod (pro určení typu útoku), remote_port (pro určení zrojové IP adresy útoku), remote_port (pro určení zdrojového portu útoku), connection (pro určení spojení), connection_root (pro určení kořene spojení) a States (pro určení země původu útoku). V každém atributu jsou obsaženy zdrojová data z databáze. Atributy jsou pro všechny analýzy útoků stejné. Po překonvertování vzniklého CSV do souboru ARFF provedla WEKA analýzu dat a připravila vizuální zpracování (viz Obrázek 4.6).



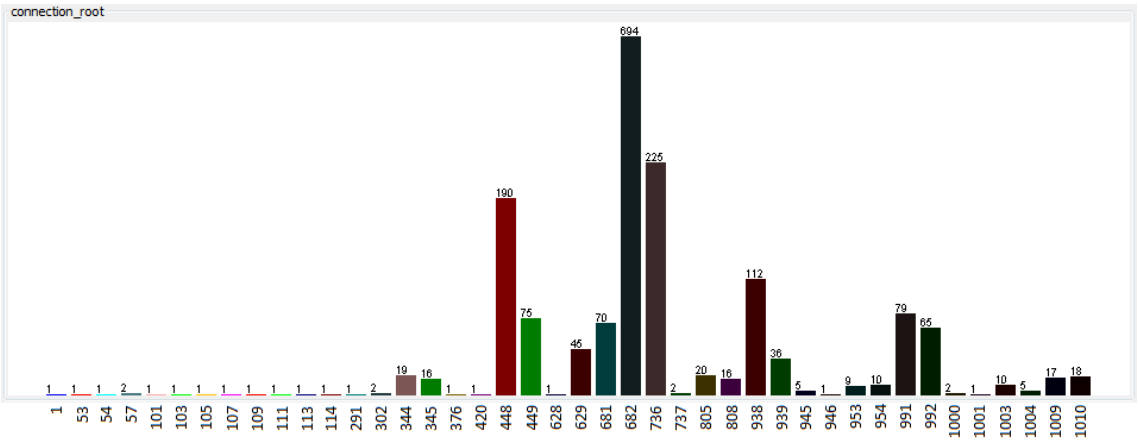
Obrázek 4.6: Atributy REGISTER v nástroji WEKA

Pro REGISTER byla použita stejná metoda jako pro analýzu útoku INVITE. Útok na REGISTER je nejčastější, zvětšilo se množství analyzovaných dat. V Příloze C jsou i další analýzy, jako státy odkud byli útoky vedeny, a také rozhodovací strom za pomoci použití algoritmu J48. Obrázek 4.7 ukazuje počty útoků, které byly provedeny.



Obrázek 4.7: IP adresy provádějící útok na REGISTER

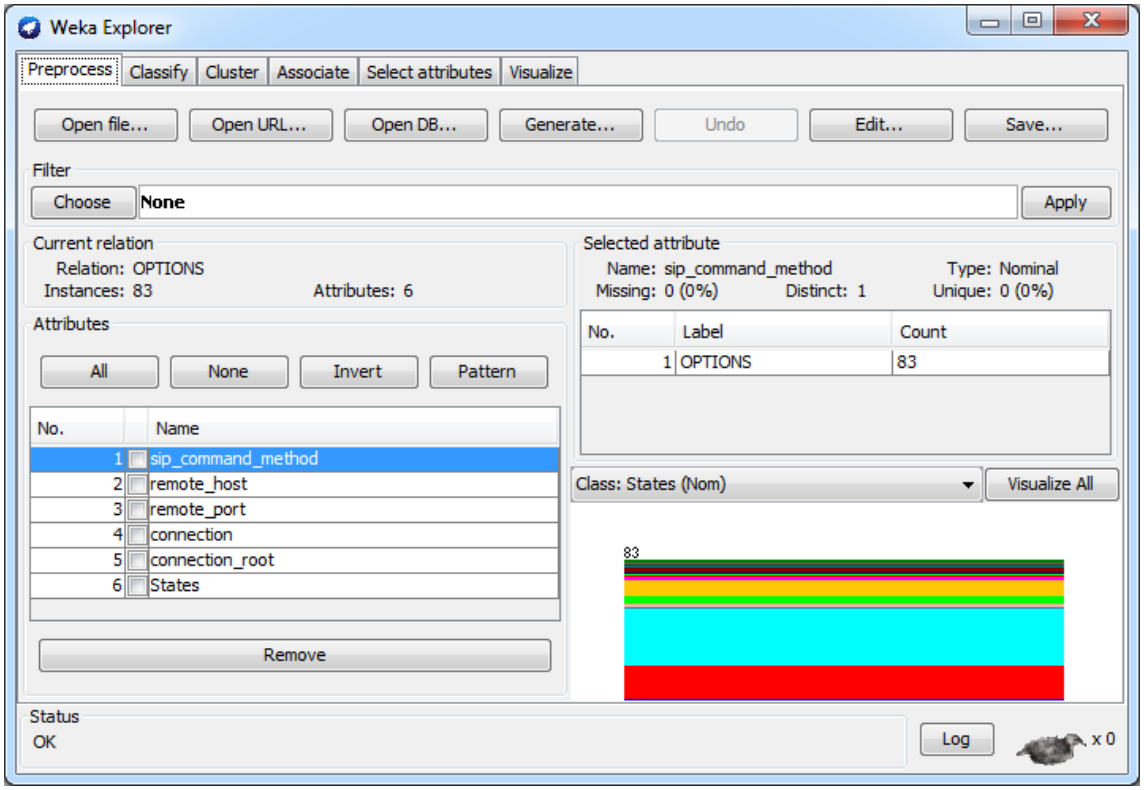
Na Obrázku 4.8 jsou zobrazena kořenová spojení a počet útoků, které byli provedeny v daném spojení.



Obrázek 4.8: Kořenová spojení a počet provedených útoků na REGISTER

4.1.3 Útok na OPTIONS

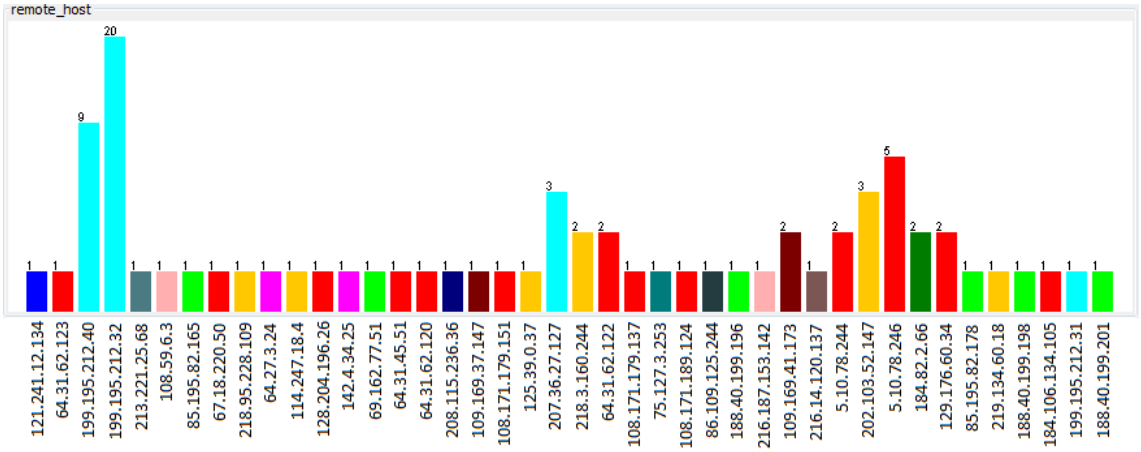
Pro analýzu OPTIONS útoku jsem si nejdříve pomocí SQL dotazu získal z databáze všechny potřebné atributy. Nastavené atributy byly SIP_command_metod (pro určení typu útoku), remote_port (pro určení zrojové IP adresy útoku), remote_port (pro určení zdrojového portu útoku), connection (pro určení spojení), connection_root (pro určení kořene spojení) a States (pro určení země původu útoku). V každém atributu jsou obsaženy zdrojová data z databáze. Atributy jsou pro všechny analýzy útoků stejné. Po překonvertování vzniklého CSV do souboru ARFF provedla WEKA analýzu dat a připravila vizuální zpracování (viz Obrázek 4.9).



Obrázek 4.9: Atributy OPTIONS v nástroji WEKA

U toho typu útoku dochází k útokům z jedné IP adresy, ale mění se port útoku, proto je každý útok zapsán jako jedno spojení v databázi.

Na Obrázku 4.10 jsou ukázány IP adresy a jejich počet, kolikrát došlo k pokusu o útok na OPTIONS.



Obrázek 4.10: IP adresy provádějící útok na REGISTER

POZNÁMKA:

Všechny výše vložené obrázky se nacházejí na přiloženém DVD v plné kvalitě.

5 Vyhodnocení úspěšnosti klasifikátoru

SIP protokoly jsou standardní pro tyto účely, a také jeden z nejvíce používaných protokolů pro manipulaci s VoIP službami. To vede k situacím, které jsou uvedeny a popsány výše v této práci. Program v Matlabu nám ohodnotí výsledek binární hodnotou a přiřadí i číslo spojení a vypíše ho do dané tabulky podle toho, o jaký útok se jednalo. Z toho vyplývá, že nám rozhodovací strom vytvoří 3 tabulky, do kterých zapisuje hodnoty i číslo jejich spojení, u typu OPTIONS vypíše jejich IP adresu. Nástroj WEKA nám udělá stejnou službu, ale s mnohem většími objemy dat a poskytne nám i mnohem větší možnosti jejich analýzy a úpravy přehledného grafického zobrazení, a také vytvoření stromu v nástroji WEKA pomocí klasifikačního algoritmu J48. Protože mé analyzované soubory měly malou velikost, byla korektnost 100%. Klasifikace, které provádí lidé jsou velmi přesná, ale časově náročná. Proto se zavádí tyto automatické klasifikátory, aby přispěly k řešení VoIP problémů. Byl proveden výzkum a byla prokázána vysoká zranitelnost těchto SIP protokolů. Tento klasifikátor je jedním ze způsobu jejich odhalení. Zlepšení bezpečnosti celé infrastruktury IP telefonie spočívá v nasazení monitorovacího mechanismu. Ten může po distribuci na monitorovacích uzlech detekovat hrozby.

Závěr

Tato bakalářská práce se zabývá problematikou bezpečnosti VoIP infrastruktury. S masivním nástupem těchto VoIP technologií se objevuje riziko způsobené především nedostatkem bezpečnostních prvků v tomto systému. Proto se zabezpečení systému stává nezbytnou nutností. Dále se tato práce zabývá analýzou dat útoků v IP telefonii za pomoci honeypotů a vytvoření klasifikátoru, který rozpozná útoky na reálných datech.

V teoretické části se zaměřuji na bezpečnostní hrozby dle VOIPSA, který popisuje a klasifikuje různé druhy VoIP útoků. Tato část popisuje, že cílem těchto útoků mohou být síťová zařízení, servery i jejich operační systémy, síťové protokoly, ale také IP telefony a jejich software. Teoretická část dále pokračuje popisem honeypotů a jejich funkcemi.

V praktické části jsem se zaměřil na rozpoznání útoků v SIP protokolu, především na příkazy − INVITE, REGISTER, OPTIONS. Nejdříve jsem si musel upravit databázi pouze pro své účely. K tomu jsem použil SQL příkazy a z celkové databáze jsem vygeneroval jednotlivé typy útoků, které měly přesně dané parametry. Po získání těchto dat jsem provedl sloučení zjednodušených databází do jediné, která byla použita jako vstupní matice pro program Matlab. V programu Matlab jsem vytvořil klasifikátor útoků v podobě rozhodovacího stromu, který rozpoznává typ útoků a binárně vyhodnotí, zda se jednalo o útok, či nikoli. Dále jsem se v praktické části zabýval analýzou útoků za pomoci nástroje WEKA. Byly vytvořeny relace pro jednotlivé typy útoků, které měly přesné dané atributy. Pro každý typ útoků (INVITE, REGISTER, OPTIONS), jsem provedl zvlášť analýzu. Použil jsem také algoritmus J48 pro vytvoření rozhodovacího stromu v tomto nástroji.

Použitá literatura

[1] DOČKAL, JAROSLAV, RICHARD MALINA, JIŘÍ MARKL a TOMÁŠ VANĚK. Bezpečnost internetové telefonie. [online]. 2006 [cit. 2014-04-30]. Dostupné z: http://www.nextsoft.cz/~malina/cs/articles/voip/clanek_Bezpecnost.pdf

[2] ENDLER, DAVID a COLLIER. Hacking Exposed VoIP: Voice Over IP Security Secrets & Solutions [online]. 2006 [cit. 2014-04-30]. ISBN 9780072263640. Dostupné z: http://www.amazon.com/Hacking-Exposed-VoIP-Security-Solutions/dp/0072263644

[3] HONEYPOT. 4SAFETY [online]. 2011 [cit. 2014-04-30]. Dostupné z: http://www.4safety.cz/text/honey

[4] ZERO.DAY EXPLOIT. SEARCHSECURITY [online]. 2010 [cit. 2014-04-30]. Dostupné z: http://searchsecurity.techtarget.com/definition/zero-day-exploit

[5] ŠAFAŘÍK, JAKUB, FILIP ŘEZÁC a MIROSLAV VOZŇÁK. Monitoring of Malicious Traffic in IP Telephony Infrastructure. [online]. 2012 [cit. 2014-04-30]. Dostupné z: http://www.cesnet.cz/wp-content/uploads/2013/02/ip-telephony-malicious-traffic-monitoring.pdf

[6] ŠAFAŘÍK, JAKUB, PAVOL PARTILA, FILIP ŘEZÁC, LUKAS MACURA a MIROSLAV VOZŇÁK. Automatic Classification of Attacks on IP Telephony. [online]. 2013 [cit. 2014-04-30]. Dostupné z: http://advances.utc.sk/index.php/AEEE/article/view/899

[7] PROVOS, NIELS a THORSTEN HOLTZ. Virtual Honeypots: From Botnet Tracking to Intrusion Detection [online]. 2008 [cit. 2014-04-30]. Dostupné z: http://books.google.cz/books?id=YQmWtsqlvfMC&printsec=frontcover&dq=virtual+honeypots&hl=cs&sa=X&ei=9CthU7-CDtSu7Aa8lICICg&ved=0CDIQ6AEwAA#v=onepage&q=virtual%20honeypots&f=false

[8] HOLMES, GEOFFREY, ANDREW DONKIN a IAN WITTEN. WEKA: A Machine Learning Workbench. [online]. [cit. 2014-05-01]. Dostupné z: http://www.cs.waikato.ac.nz/~ml/publications/1994/Holmes-ANZIIS-WEKA.pdf

[9] Attribute-relation file format (arff). In: [online]. [cit. 2014-05-01]. Dostupné z: http://www.cs.waikato.ac.nz/ml/weka/arff.html

[10] VOIPSA. VoIP Security and Privacy Threat Taxonomy. [online]. 2005 [cit. 2014-05-02]. Dostupné z: http://www.voipsa.org/Activities/VOIPSA_Threat_Taxonomy_0.1.pdf

Seznam příloh

Příloha A: Rozdělené typy útoků I

Příloha B: Program rozhodovacího stromu v MATLABU VI

Příloha C: Výsledky analýzy nástroje WEKA IX

Součástí BP/DP je CD/DVD.

Adresářová struktura přiloženého CD/DVD:

1. Program pro Matlab
2. Použitá databáze
3. INVITE data
4. INVITE výsledky
5. REGISTER data
6. REGISTER výsledky
7. OPTIONS data
8. OPTIONS výsledky
9. WEKA ALL data
10. WEKA ALL výsledky
11. Použité programy

Příloha A: *Rozdělené typy útoků*

Tabulka A.1: Tabulka všech útoků typu INVITE

sip command method	remote host	remote port	connection	connection root	Počet útoků	Země útoku
INVITE	192.168.255.1	9484	2	1	1	
INVITE	192.168.255.1	9484	40	39	1	
INVITE	37.8.54.135	10083	382	376	6	Izrael
INVITE	37.8.54.135	10083	408	383	25	Izrael
INVITE	37.8.54.135	10083	419	409	10	Izrael
INVITE	87.230.26.233	5060	935	906	29	Německo
INVITE	198.101.195.124	5082	958	957	1	USA-Texas
INVITE	198.101.195.124	5088	960	959	1	USA-Texas
INVITE	198.101.195.124	5082	962	961	1	USA-Texas
INVITE	198.101.195.124	5082	964	963	1	USA-Texas
INVITE	198.101.195.124	5071	968	966	2	USA-Texas
INVITE	198.101.195.124	5078	970	969	1	USA-Texas
INVITE	198.101.195.124	5071	973	972	1	USA-Texas
INVITE	198.101.195.124	5071	975	974	1	USA-Texas
INVITE	198.101.195.124	5088	977	976	1	USA-Texas
INVITE	198.101.195.124	5076	980	979	1	USA-Texas
INVITE	198.101.195.124	5082	982	981	1	USA-Texas
INVITE	198.101.195.124	5082	984	983	1	USA-Texas
INVITE	198.101.195.124	5074	987	986	1	USA-Texas
INVITE	198.101.195.124	5071	989	988	1	USA-Texas

Tabulka A.2: Tabulka všech útoků REGISTER

sip command method	remote host	remote port	connection	connection root	Počet útoků	Země útoku
REGISTER	192.168.255.1	9484	1	1	1	
REGISTER	192.168.255.1	9486	53	53	1	
REGISTER	192.168.255.1	9488	54	54	1	
REGISTER	192.168.255.1	9488	57	57	2	
REGISTER	192.168.255.1	59696	101	101	1	
REGISTER	192.168.255.1	59700	103	103	1	
REGISTER	192.168.255.1	59701	105	105	1	
REGISTER	192.168.255.1	59710	107	107	1	
REGISTER	192.168.255.1	59714	109	109	1	
REGISTER	192.168.255.1	59715	111	111	1	
REGISTER	192.168.255.1	59716	113	113	1	
REGISTER	192.168.255.1	9489	114	114	1	
REGISTER	192.168.255.1	9490	291	291	1	
REGISTER	192.168.255.1	9490	302	302	2	
REGISTER	64.31.62.123	5801	344	344	19	USA-Texas
REGISTER	64.31.62.123	5189	345	345	16	USA-Texas
REGISTER	37.8.54.135	10083	376	376	1	Izrael
REGISTER	37.8.54.135	10083	420	420	1	Izrael
REGISTER	218.95.228.109	5080	448	448	190	Čína
REGISTER	218.95.228.109	5088	449	449	75	Čína
REGISTER	64.31.62.120	5918	628	628	1	USA-Texas
REGISTER	64.31.62.120	5528	629	629	45	USA-Texas
REGISTER	109.169.37.147	5075	681	681	70	Anglie
REGISTER	109.169.37.147	5077	682	682	694	Anglie
REGISTER	125.39.0.37	5100	736	736	225	Čína
REGISTER	125.39.0.37	5103	737	737	2	Čína
REGISTER	64.31.62.122	5429	805	805	20	USA-Texas
REGISTER	64.31.62.122	5715	808	808	16	USA-Texas
REGISTER	216.187.153.142	5086	938	938	112	USA-Virginie
REGISTER	216.187.153.142	5087	939	939	36	USA-Virginie
REGISTER	5.10.78.244	5521	945	945	5	USA-Texas
REGISTER	5.10.78.244	6001	946	946	1	USA-Texas
REGISTER	5.10.78.246	5933	953	953	9	USA-Texas
REGISTER	5.10.78.246	5967	954	954	10	USA-Texas
REGISTER	64.31.62.122	5383	991	991	79	USA-Texas
REGISTER	64.31.62.122	5143	992	992	65	USA-Texas
REGISTER	5.10.78.246	5107	1000	1000	2	USA-Texas
REGISTER	5.10.78.246	5078	1001	1001	1	USA-Texas
REGISTER	188.40.199.198	5387	1003	1003	10	Německo
REGISTER	188.40.199.198	5127	1004	1004	5	Německo
REGISTER	188.40.199.201	5151	1009	1009	17	Německo
REGISTER	188.40.199.201	5164	1010	1010	18	Německo

Tabulka A.3: Tabulka všech útoků typu OPTIONS

sip command method	remote host	remote port	connection	connection root	Počet útoků	Země útoku
OPTIONS	121.241.12.134	5062	323	323	1	Indie
OPTIONS	64.31.62.123	5179	343	343	1	USA-Texas
OPTIONS	199.195.212.40	5068	351	351	1	USA-Florida
OPTIONS	199.195.212.40	5060	352	352	1	USA-Florida
OPTIONS	199.195.212.40	5076	353	353	1	USA-Florida
OPTIONS	199.195.212.40	5081	355	355	1	USA-Florida
OPTIONS	199.195.212.40	5079	356	356	1	USA-Florida
OPTIONS	199.195.212.40	5070	357	357	1	USA-Florida
OPTIONS	199.195.212.40	5088	358	358	1	USA-Florida
OPTIONS	199.195.212.40	5069	359	359	1	USA-Florida
OPTIONS	199.195.212.40	5066	365	365	1	USA-Florida
OPTIONS	199.195.212.32	5063	425	425	1	USA-Florida
OPTIONS	213.221.25.68	5065	434	434	1	Rusko
OPTIONS	108.59.6.3	5078	436	436	1	USA-Virinie
OPTIONS	85.195.82.165	5060	443	443	1	Německo
OPTIONS	67.18.220.50	5060	445	445	1	USA-Texas
OPTIONS	218.95.228.109	5071	447	447	1	Čína
OPTIONS	64.27.3.24	5088	453	453	1	USA-Kalifornie
OPTIONS	114.247.18.4	5060	454	454	1	Čína
OPTIONS	199.195.212.32	5130	456	456	1	USA-Florida
OPTIONS	199.195.212.32	5071	525	525	1	USA-Florida
OPTIONS	199.195.212.32	5096	550	550	1	USA-Florida
OPTIONS	128.204.196.26	5064	575	575	1	Nizozemsko
OPTIONS	142.4.34.25	5060	587	587	1	USA-Kalifornie
OPTIONS	69.162.77.51	5200	592	592	1	USA-Indiana
OPTIONS	199.195.212.32	5086	594	594	1	USA-Florida
OPTIONS	64.31.45.51	5063	605	605	1	USA-Texas
OPTIONS	64.31.62.120	5223	625	625	1	USA-Texas
OPTIONS	208.115.236.36	5068	634	634	1	USA-Georgia
OPTIONS	199.195.212.32	5134	665	665	1	USA-Florida
OPTIONS	109.169.37.147	5060	680	680	1	Anglie
OPTIONS	108.171.179.151	5068	693	693	1	USA-Texas
OPTIONS	125.39.0.37	5060	735	735	1	Čína
OPTIONS	207.36.27.127	5070	743	743	1	USA-Florida
OPTIONS	199.195.212.32	5065	762	762	1	USA-Florida
OPTIONS	218.3.160.244	5074	769	769	1	Čína
OPTIONS	199.195.212.32	5080	771	771	1	USA-Florida
OPTIONS	64.31.62.122	5351	804	804	1	USA-Texas
OPTIONS	108.171.179.137	5103	810	810	1	USA-Texas
OPTIONS	75.127.3.253	5065	825	825	1	USA-New York
OPTIONS	199.195.212.32	5066	829	829	1	USA-Florida
OPTIONS	199.195.212.32	5077	832	832	1	USA-Florida
OPTIONS	108.171.189.124	5061	835	835	1	USA-Texas
OPTIONS	207.36.27.127	5062	859	859	1	USA-Florida
OPTIONS	199.195.212.32	5063	902	902	1	USA-Florida
OPTIONS	86.109.125.244	5135	903	903	1	Španělsko
OPTIONS	188.40.199.196	5102	904	904	1	Německo
OPTIONS	199.195.212.32	5096	905	905	1	USA-Florida
OPTIONS	216.187.153.142	5069	937	937	1	USA-Virinie

(pokračování tabulky na straně IV)

(pokračování tabulky ze strany III)

sip command method	remote host	remote port	connection	connection root	Počet útoků	Země útoku
OPTIONS	109.169.41.173	5062	940	940	1	Anglie
OPTIONS	216.14.120.137	5060	941	941	1	USA-Ilinois
OPTIONS	109.169.41.173	5062	942	942	1	Anglie
OPTIONS	199.195.212.32	5130	943	943	1	USA-Florida
OPTIONS	5.10.78.244	5153	944	944	1	USA-Texas
OPTIONS	202.103.52.147	5060	947	947	1	Čína
OPTIONS	199.195.212.32	5075	948	948	1	USA-Florida
OPTIONS	5.10.78.246	5094	949	949	1	USA-Texas
OPTIONS	184.82.2.66	5060	950	950	1	USA-Pensilvanie
OPTIONS	129.176.60.34	5066	951	951	1	USA-Texas
OPTIONS	5.10.78.246	5062	952	952	1	USA-Texas
OPTIONS	5.10.78.244	5153	944	944	1	USA-Texas
OPTIONS	202.103.52.147	5060	947	947	1	Čína
OPTIONS	199.195.212.32	5075	948	948	1	USA-Florida
OPTIONS	5.10.78.246	5094	949	949	1	USA-Texas
OPTIONS	184.82.2.66	5060	950	950	1	USA-Pensilvanie
OPTIONS	129.176.60.34	5066	951	951	1	USA-Texas
OPTIONS	5.10.78.246	5062	952	952	1	USA-Texas
OPTIONS	207.36.27.127	5079	955	955	1	USA-Florida
OPTIONS	199.195.212.32	5130	956	956	1	USA-Florida
OPTIONS	64.31.62.122	5200	990	990	1	USA-Texas
OPTIONS	199.195.212.32	5078	993	993	1	USA-Florida
OPTIONS	218.3.160.244	5091	994	994	1	Čína
OPTIONS	85.195.82.178	5062	995	995	1	Německo
OPTIONS	199.195.212.32	5104	996	996	1	USA-Florida
OPTIONS	199.195.212.32	5106	997	997	1	USA-Florida
OPTIONS	219.134.60.18	5075	998	998	1	Čína
OPTIONS	5.10.78.246	5087	999	999	1	USA-Texas
OPTIONS	188.40.199.198	5102	1002	1002	1	Německo
OPTIONS	184.106.134.105	5060	1005	1005	1	USA-Texas
OPTIONS	202.103.52.147	5060	1006	1006	1	Čína
OPTIONS	199.195.212.31	5075	1007	1007	1	USA-Florida
OPTIONS	188.40.199.201	5131	1008	1008	1	Německo
OPTIONS	199.195.212.32	5156	1011	1011	1	USA-Florida

Tabulka A.4: Tabulka všech útoků typu OPTIONS seřazena dle IP

sip command method	remote host	Počet útoků	sip command method	remote host	Počet útoků
OPTIONS	108.171.179.137	1	OPTIONS	207.36.27.127	3
OPTIONS	108.171.179.151	1	OPTIONS	208.115.236.36	1
OPTIONS	108.171.189.124	1	OPTIONS	213.221.25.68	1
OPTIONS	108.59.6.3	1	OPTIONS	216.14.120.137	1
OPTIONS	109.169.37.147	1	OPTIONS	216.187.153.142	1
OPTIONS	109.169.41.173	2	OPTIONS	218.3.160.244	2
OPTIONS	114.247.18.4	1	OPTIONS	218.95.228.109	1
OPTIONS	121.241.12.134	1	OPTIONS	219.134.60.18	1
OPTIONS	125.39.0.37	1	OPTIONS	5.10.78.244	1
OPTIONS	128.204.196.26	1	OPTIONS	5.10.78.246	3
OPTIONS	129.176.60.34	1	OPTIONS	64.27.3.24	1
OPTIONS	142.4.34.25	1	OPTIONS	64.31.45.51	1
OPTIONS	184.106.134.105	1	OPTIONS	64.31.62.120	1
OPTIONS	184.82.2.66	1	OPTIONS	64.31.62.122	2
OPTIONS	188.40.199.196	1	OPTIONS	64.31.62.123	1
OPTIONS	188.40.199.198	1	OPTIONS	67.18.220.50	1
OPTIONS	188.40.199.201	1	OPTIONS	69.162.77.51	1
OPTIONS	199.195.212.31	1	OPTIONS	75.127.3.253	1
OPTIONS	199.195.212.32	19	OPTIONS	85.195.82.165	1
OPTIONS	199.195.212.40	9	OPTIONS	85.195.82.178	1
OPTIONS	202.103.52.147	2	OPTIONS	86.109.125.244	1

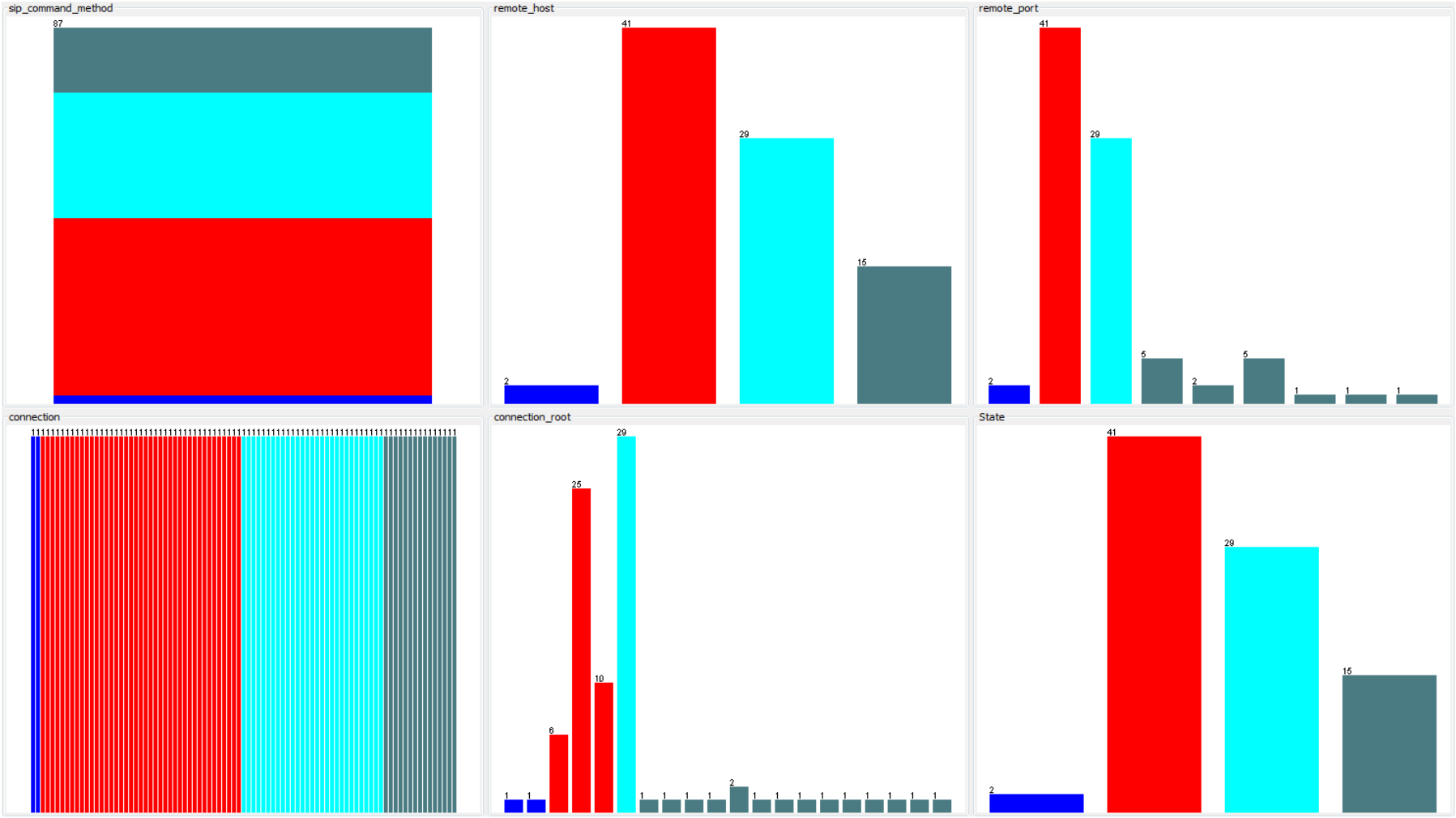
Příloha B: *Program rozhodovacího stromu v MATLABU*

```
M = csvread('souhrn.csv'); %načtení cvs souboru obsahující
numericko hodnoty oddělené čárkou do matice M
[r, c] = size(M); %velikost matice M

%zjištěni počtu proměnných INVITE, REGISTER, OPTIONS
cin = 0; %počet proměnných INVITE
cre = 0;
cot = 0;
for x = 1:r

    if M(x, 1) == 1 %pokud je proměnná INVITE
        cin = cin + 1; %zvyš čítač cin
    elseif M(x, 1) == 2 %pokud je proměnná REGISTER
       cre = cre + 1; %zvyš čítač cre
    else
      cot = cot + 1; %zvyš čítač cot
    end
end

atin = zeros(cin, 1); %vektor pro INVITE, ve kterém je uloženo,
zda se jednalo o utok (1) nebo normální komunikaci(0)
atre = zeros(cre, 1);
atot = zeros(cot, 1);

pin = 1; %aktuální ukazatel na pozici v poli atin
pre = 1;
pot = 1;

%rozhodovaci strom
for x = 1:r
```

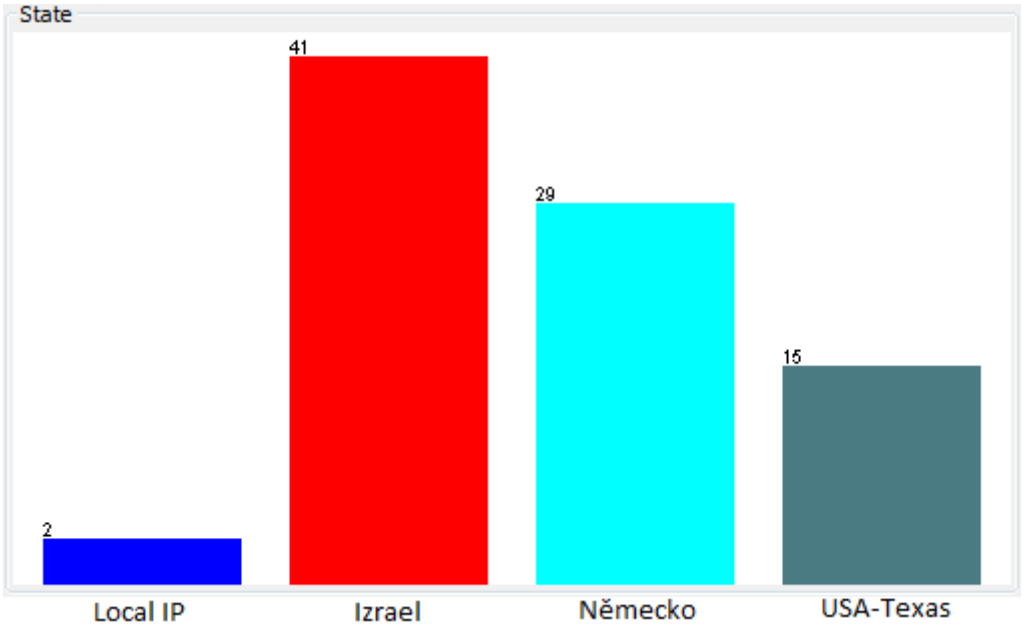
```
    if M(x, 1) == 3
        tmp = M(x, c);
    else
        tmp = M(x, 3);
    end

    if tmp > 3
       if M(x, 1) == 1
          atin(pin) = 1;
          pin = pin + 1;
       elseif M(x, 1) == 2
           atre(pre) = 1;
           pre = pre + 1;
       else
           atot(pot) = 1;
          pot = pot + 1;
       end
    else
       if M(x, 1) == 1
          atin(pin) = 0;
          pin = pin + 1;
       elseif M(x, 1) == 2
           atre(pre) = 0;
           pre = pre + 1;
       else
           atot(pot) = 0;
           pot = pot + 1;
       end
    end
end
```

```
display( 'INVITE:' );

for x = 1:cin
    display( {atin(x), M(x, 2) } );
end

display( 'REGISTER:' );

for x = 1:cre
    display( {atre(x), M(x + cin, 2) } );
end

display( 'OPTIONS:' );

for x = 1:cot
    ip = '';

    for y = 2:4
        ip = strcat( ip, int2str( M(x + cre + cin, y) ), '.' );
    end
    ip = strcat( ip, int2str( M(x + cre + cin, 5) ) );

    display( { atot(x), ip } );
end
```

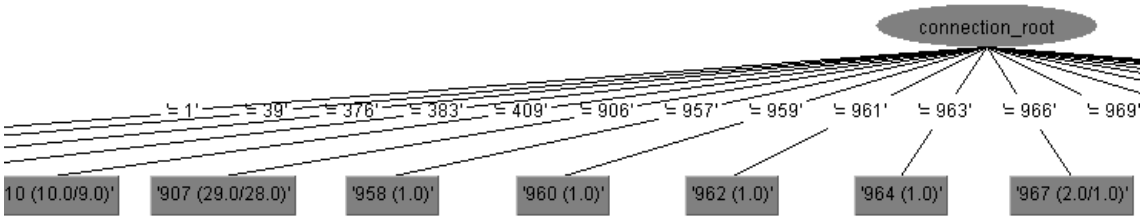
Příloha C: *Výsledky analýzy nástroje WEKA*



Obrázek C.1: Souhrná vizualizace výsledků pro útok na INVITE

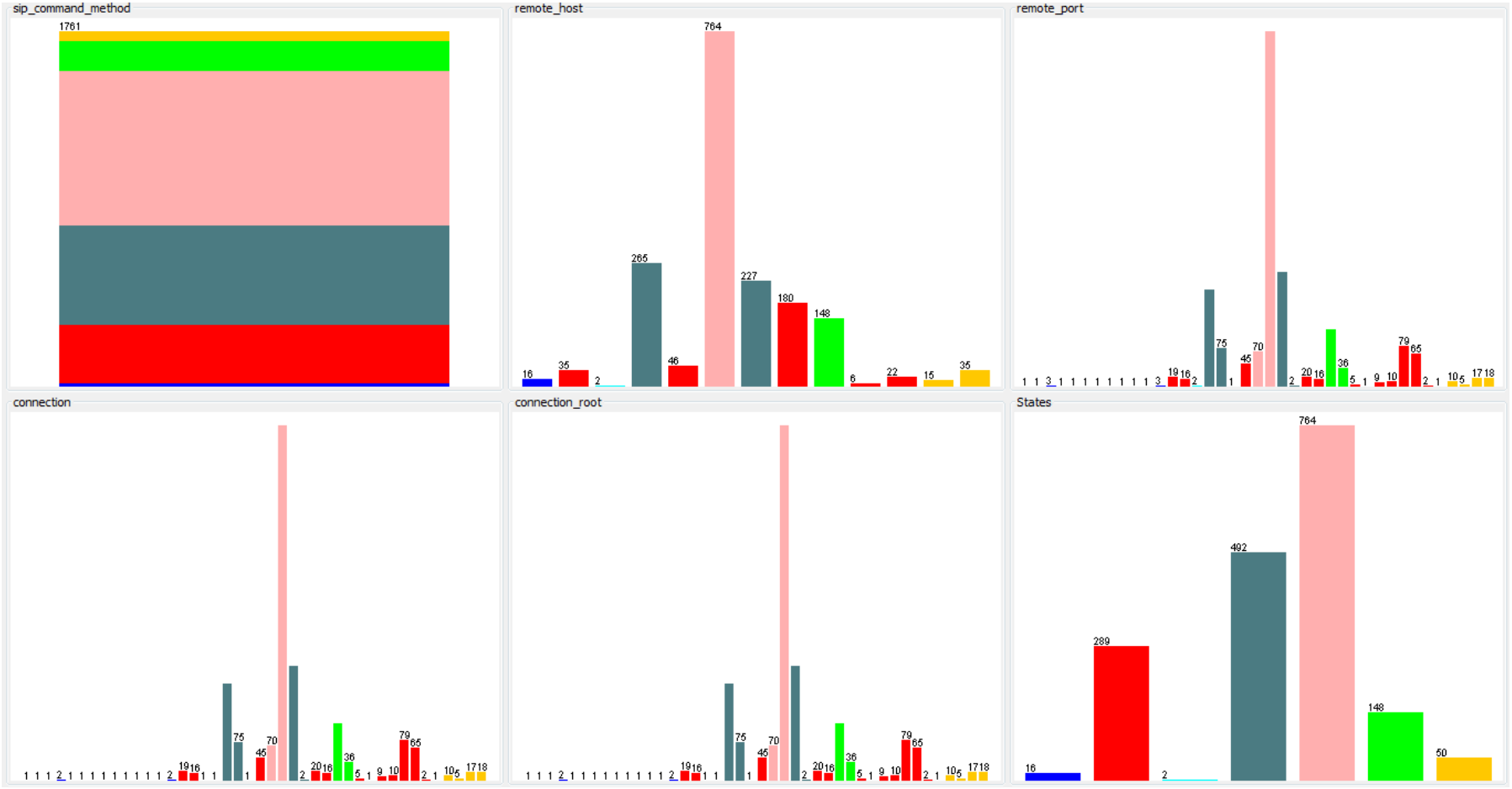


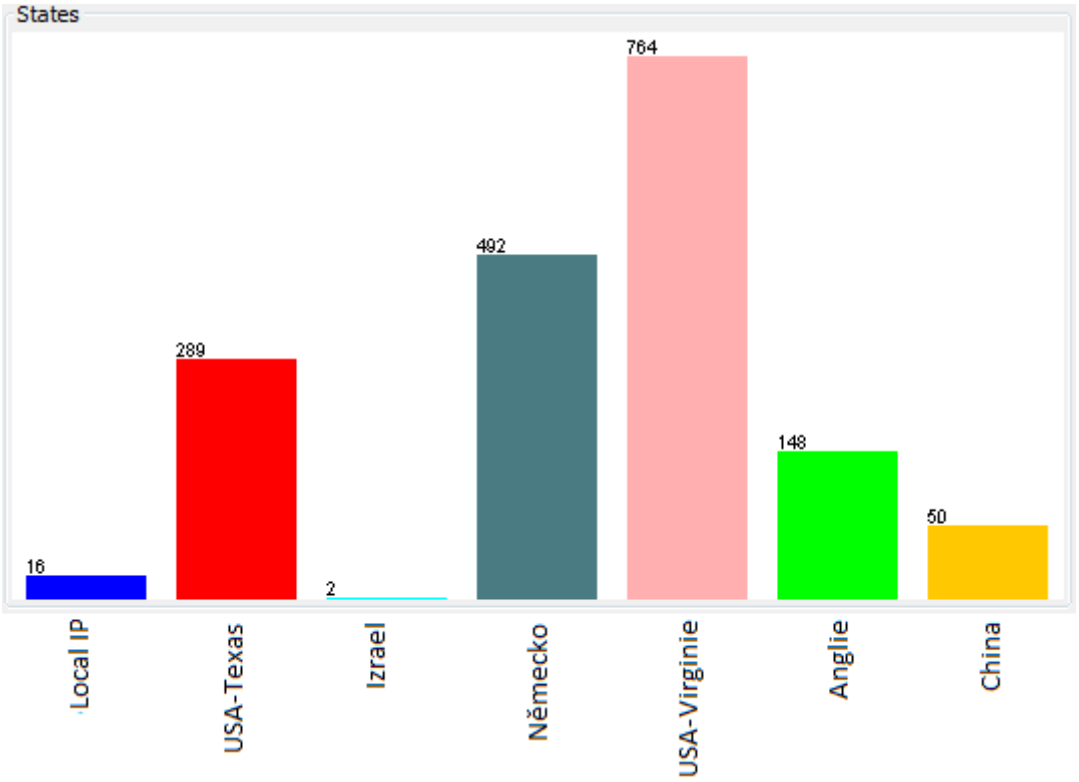
Obrázek C.2: Země původu útoku na INVITE



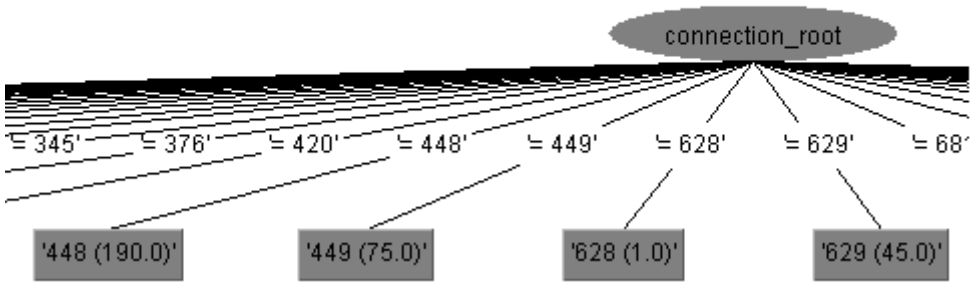
Obrázek C.3: Výřez rozhodovacího stromu J48 pro INVITE z WEKY

Obrázek C.4: *Souhrnná vizualizace výsledků pro REGISTER*



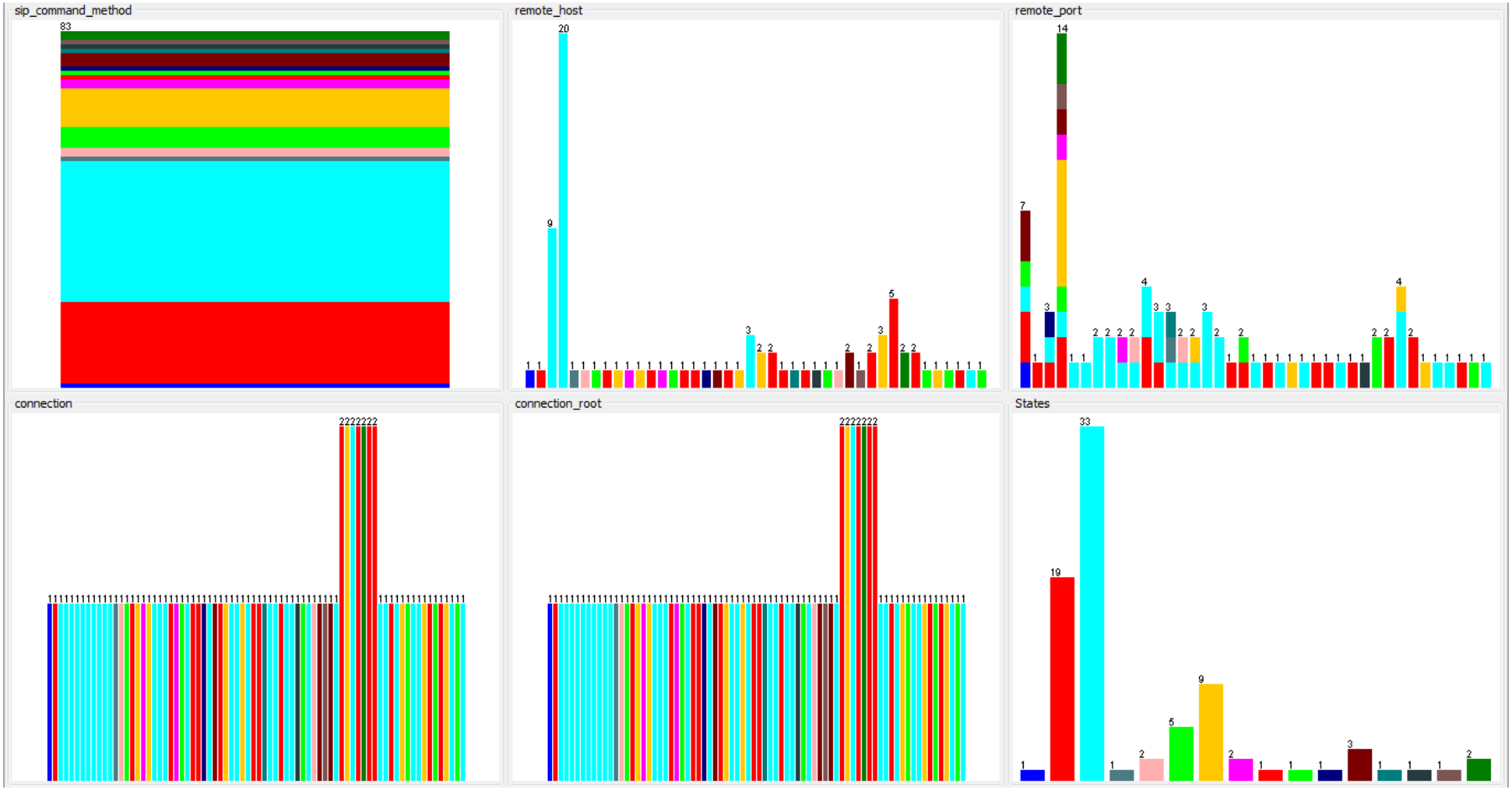


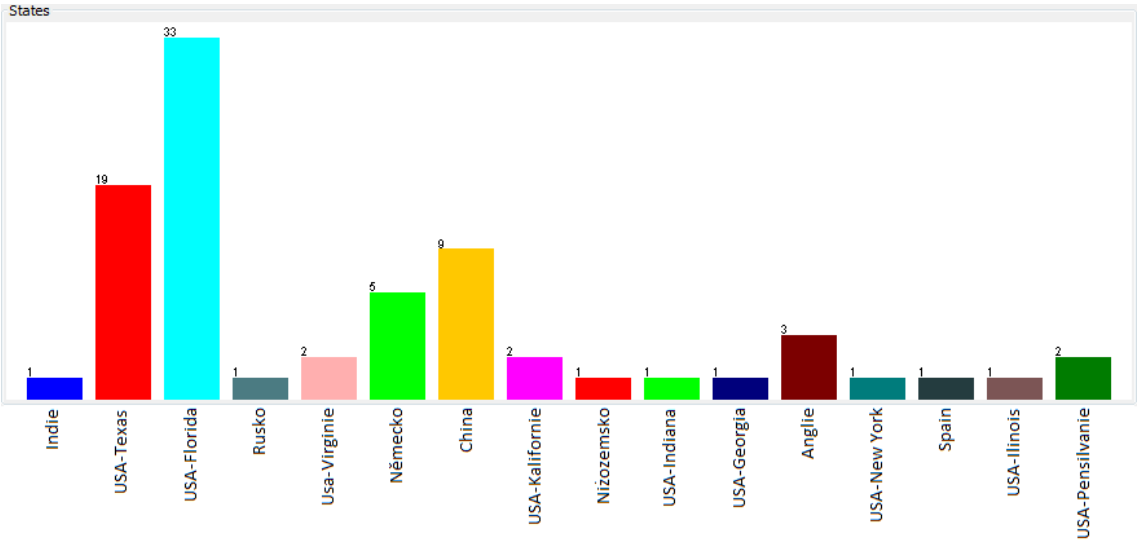
Obrázek C.5: *Země původu útok na REGISTER*



Obrázek C.6: Výřez rozhodovacího stromu J48 pro REGISTER z WEKY

Obrázek C.7: *Souhrnná vizualice výledů pro OPTIONS*





Obrázek C.8: Země původu útok na OPTIONS

Obrázek C.9: *Souhrnná vizualizace všech tří útoků současně*